Enhancing dental radiographic interpretation by collaborating with AI systems to minimise interpretive errors

Shwetha Hegde

Sydney Dental School

Faculty of Medicine and Health

The University of Sydney



A thesis submitted to fulfil requirements for the degree of Doctor of Philosophy

The research reported in this thesis was supported by the award of a Research Training Program scholarship to the PhD Candidate

2025

Table of Contents

Abstract....4

Statement of originality....6

Publications during the candidature....7

Authorship attribution statement....8

Conference presentations....9

List of figures and tables....10

Acknowledgements....11

1. Introduction

1.1 Structure of the thesis.... 12

1.2 Motivation for the thesis....12

1.3 Research objectives....16

1.4 Grants....17

1.5 References....17

2.Factors Affecting the Interpretation of Dental Radiographs - Systematic Review....21

3. Australian Dentist's Knowledge and Perceptions of Factors Affecting Radiographic Interpretation....34

4. Australian Dentist's Knowledge of Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures....53

5. Machine Learning Algorithms Enhance the Accuracy of Radiographic Diagnosis of Dental Caries: A Comparative Study....81

6. Attitudes and Perceptions of Australian Dentists and Dental Students Towards Applications of Artificial Intelligence in Dentistry: A survey....130

7.Machinometrics, a novel theoretical framework for explainability in healthcare....159

8. General Discussion and Conclusions

8.1 Analysis of Factors Contributing to Interpretive Errors....220

8.2 Role of Cognitive Aids in Diagnostic Accuracy....221

8.3 AI in Dentistry: Opportunities and Challenges....222

8.4 Summary of Key Findings....223

8.5 Recommendations for Future Research.....224

8.6 References.....224

9. Appendices

9.1 Ethics approval letter (chapters 3 and 4).....227

9.2 Ethics approval letter (chapter 6).....229

9.3 Ethics approval letter (chapter 5).....231

Abstract

Interpretive errors in radiology, particularly in dental radiography, pose significant risks to patient care due to their potential impact on diagnosis and treatment.[1] These errors often stem from limitations in human capabilities, including visual detection, pattern recognition, and clinical reasoning.[2] Despite the critical role of radiographic assessments in clinical dentistry, there is a notable lack of evidence regarding the prevalence and causes of errors in dental radiograph interpretation. This PhD thesis addresses this gap by evaluating factors influencing these errors and examining the effectiveness of machine learning (ML) algorithms as tools to minimise them.

The study involved a systematic review,[3] surveys of dental practitioners[4,5] and students[6] in Australia, and a comparative study on using cognitive aids, such as ML algorithms and checklists, in diagnosing dental caries from bitewing radiographs.[7] Key findings highlighted cognitive overload, poor image quality, lack of experience, and excessive workload as significant contributors to interpretive errors. Errors of omission were noted as the most common error type, with direct consequences for patient and clinician-related, including undertreatment (72%), increased costs (62%), legal issues (82%), and reputational damage (75.6%).

ML algorithms were found to improve the accuracy of caries diagnosis, with higher sensitivity rates (79%) ($p<0.001$) compared to checklist and control groups. The confidence level in diagnosis among participants in the ML group was also higher than that of traditional methods. The potential of AI in dentistry was recognised, especially in areas like dentomaxillofacial radiology, implantology, and prosthodontics, despite concerns over job loss, accuracy and trust in AI outputs. A novel explainability method was developed to allow clinicians and computer scientists to understand ML explanations easily.[8]

In conclusion, the thesis highlights the role of assistive cognitive aids, particularly ML algorithms, in reducing interpretive errors in dental radiology, suggesting a supportive pathway for enhancing diagnostic accuracy and clinician confidence in dental care.

References:

1. Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual and Interpretive Error in Diagnostic Radiology—Causes and Potential Solutions. Acad Radiol [Internet]. 2019;26(6):833–45. Available from: https://www.sciencedirect.com/science/article/pii/S107663321830521X

2. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive Error in Radiology. AJR Am J Roentgenol. 2017 Apr;208(4):739–49.

3. Hegde S, Gao J, Vasa R, Cox S. Factors affecting interpretation of dental radiographs. Dentomaxillofacial Radiol [Internet]. 2022 Dec 22;52(2):20220279. Available from: https://doi.org/10.1259/dmfr.20220279

4. Hegde S, Gao J, Nanayakkara S, Vasa R, Cox S. Australian Dentist's Knowledge and Perceptions of Factors Affecting Radiographic Interpretation. Int Dent J. 2024; 589-596

5. Hegde S, Nanayakkara S, Cox S, Vasa R, Gao J Australian Dentists Knowledge of Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures Clinical and Experimental Dental Research 2024 Dec;10(6):e70027. https://doi: 10.1002/cre2.70027

6. Hegde S et al Attitudes And Perceptions of Australian Dental Practitioners and Dental Students Towards Applications Of Artificial Intelligence In Dentistry: A Survey European Journal of Dental Education (2024) https://doi.org/10.1111/eje.13042

7. Hegde S et al Influence Of Cognitive Aids On The Radiographic Diagnosis Of Dental Caries: A Comparative Study *(manuscript under draft)*

8. Hegde S et al Machinometrics, a new framework towards explainable machine learning in healthcare *(under revisions with BMC Methods)*

Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my work and has not been submitted for any other degree.

I certify that the intellectual content of this thesis is a result of my own work, and I have acknowledged the assistance received while preparing this thesis.

Shwetha Hegde

October, 2024

Publications during the candidature

	Title	Journal	Status	Co-authors
1	Factors affecting interpretation of dental radiographs	Dentomaxillofacial Radiology(2022) 51, 20220279 https://doi.org/10.1259/dmfr.20220279	Published	Shwetha Hegde (first and corresponding author), Jinlong Gao, Rajesh Vasa, Stephen Cox
2	Australian dentist's knowledge and Perceptions of factors affecting radiographic interpretation	International Dental Journal 74(2024) 589-596 https://doi.org/10.1016/j.identj.2023.11.006	Published	Shwetha Hegde (first and corresponding author), Jinlong Gao, Rajesh Vasa, Shanika Nanayakkara, Stephen Cox
3	Australian Dentist's Knowledge of Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures	Clinical and Experimental Dental Research (2024) DOI: 10.1002/cre2.70027	Published	Shwetha Hegde (first and corresponding author), Shanika Nanayakkara, Rajesh Vasa, Stephen Cox, Jinlong Gao
4	Machine learning algorithms enhance the accuracy of radiographic diagnosis of dental caries: A comparative study	-	Manuscript under preparation	Shwetha Hegde (first and corresponding author), Jinlong Gao, Stephen Cox, Shanika Nanayakkara, Rajesh Vasa
5	Attitudes And Perceptions of Australian Dentists and Dental Students Towards Applications of Artificial Intelligence In Dentistry: A Survey	European Journal of Dental Education (2024) https://doi.org/10.1111/eje.13042	Published	Shwetha Hegde (first and corresponding author), Shanika Nanayakkara, Ashleigh Jordan, Omar Jeha, Usaamah Patel, Vivian Luu, Jinlong Gao
6	Machinometrics, a new framework towards explainable machine learning in healthcare	BMC Methods	Under revisions	Jionghui Lin, Shwetha Hegde (co-first author), Daochang Liu, Ka Wing Cheng, Heiko Spallek, Chang Xu, Jinlong Gao

Authorship attribution statement

This thesis contains works published as:

1. **Hegde S,** Gao J, Vasa R, Cox S. Factors affecting interpretation of dental radiographs. Dentomaxillofacial Radiol. 2023 Jan;52(2):20220279. doi: 10.1259/dmfr.20220279. Epub 2022 Dec 22. PMID: 36472942; PMCID: PMC9974235. (Q1, SJR 0.87)
2. **Hegde S,** Gao J, Vasa R, Nanayakkara S, Cox S. Australian dentist's knowledge and perceptions of factors affecting radiographic interpretation International Dental Journal 2024 Jan https://doi.org/10.1016/j.identj.2023.11.006 (Q1, SJR 0.73)
3. **Hegde S,** Vasa R, Nanayakkara S, Cox S, Gao J Australian Dentist's Knowledge of Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures. Clinical and Experimental Dental Journal 2024 https://doi: 10.1002/ cre2.70027(Q2, SJR 0.54)
4. **Hegde S** et al., Machine learning algorithms enhance the accuracy of radiographic diagnosis of dental caries: A comparative study (manuscript under preparation)
5. **Hegde S,** Nanayakkara S, Jordan A, Jeha O, Patel U, Luu V, Gao J. Attitudes and Perceptions of Australian Dentists and Dental Students Towards Applications of Artificial Intelligence in Dentistry: A Survey. Eur J Dent Educ. 2024 Sep 28. doi: 10.1111/eje.13042. Epub ahead of print. PMID: 39340812. (Q1, SJR 0.63)
6. **Hegde S** et al., Machinometrics, a new framework towards explainable machine learning in healthcare. (Under revisions with BMC Methods)

Chapters 2,3,4,5,6 and 7 are listed as papers 1,2,3,4,5 and 6, respectively.

As the first author of the above articles, I designed and conducted the studies and wrote the original drafts of the manuscripts.

Shwetha Hegde

As the supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Associate Professor Stephen Cox

Conference presentations

1. 'Errors in the interpretation of dental radiographs' presented at Early to Mid Career Researcher breakfast seminar, held at Charles Perkins Center, University of Sydney on 1st of December 2022
2. 'Survey of dental practitioners in NSW, Australia, about the knowledge and perceptions of errors in interpretation of dental radiographs' pitch presentation (online) at IADMFR conference 6-8th July 2023
3. 'Attitudes and perceptions of Australian dentists and dental Students towards applications of artificial intelligence in dentistry: A survey' presented at the International Association of Dental, Oral and Craniofacial Research Australia New Zealand Conference 2023, held at Sydney, Australia 27-29th September 2023
4. 'A comparative study on the effectiveness of cognitive aids in the radiographic diagnosis of dental caries' presented at the Faculty of Medicine and Health Higher Degree by Research Conference held at Charles Perkins Center, University of Sydney, Australia on 24-25th July 2024

List of Figures and Tables

Figure 1. Concept map depicting the trajectory of the thesis.....15

Figure 2. Diagrammatic representation of the types of thinking used in radiographic analysis.....13

Figure 3. Diagrammatic representation of the role of ML as a cognitive aid.....208

Acknowledgements

Pursuing a PhD has been a personal aspiration motivated by my passion for advancing dental radiology and improving patient care. This journey has been challenging and rewarding, allowing me to grow as a researcher and clinician.

I would like to extend my heartfelt gratitude to my supervisors, Associate Professor Stephen Cox, Associate Professor Jinlong Gao and Professor Rajesh Vasa, for your invaluable experience, advice, guidance and collegiality throughout this process. Your expertise and unwavering support have been instrumental in the completion of this thesis.

To my mum and dad, Pushpa and Karunakar, thank you for your constant encouragement and belief in me. Your support has been the cornerstone of my life, providing me with the confidence to pursue my goals.

A special thanks to my husband, Sathya, for your steadfast support, the countless walks and talks, and for being a source of constant encouragement, love, and understanding. Your companionship has made this journey not only possible but also enriching. To my daughter, Anvi, your sense of humour and kindness have brought immense joy. Thank you for reminding me of the importance of balance in life and for always brightening my days.

I am also grateful to my friends and colleagues for their support and encouragement, which have been valuable in my personal and professional growth.

Shwetha Hegde

October, 2024

Chapter 1

Introduction

1.1 Structure of thesis:

The body of research presented in this thesis examines how interpretive errors can be minimised by collaborating with machine learning algorithms. Presented as a 'thesis including publications', it comprises six chapters corresponding to publications that resulted from this research. This introductory narrative, together with the concept map (Figure 1), draws the six papers together to describe the trajectory of the overall thesis. This introduction also outlines the motivation for the thesis and describes the research objectives.

The thesis contains six papers presented in chapters 2 to 7. Of these, four are published, one is under review with a journal, and one manuscript is under preparation. The general discussion and conclusion chapter examines the outcomes of this research, summarises the key findings, and identifies the implications and future directions of this work.

1.2 Motivation for the thesis:

Radiographs and imaging play a crucial role in the diagnostic process, providing essential information that guides clinical decisions. They are necessary for diagnosing and treating various dental conditions, from dental caries and periodontal disease to periapical, developmental, and other abnormalities.[1] The evolution from traditional film-based radiography to digital imaging has significantly enhanced diagnostic accuracy and reduced radiation dose to the patient. With imaging techniques offering three-dimensional views, better planning of complex dental procedures such as implant placement and orthodontic interventions is possible. [2,3] Interpretation of radiographs involves differentiating normal from abnormal findings and correlating these findings with clinical symptoms.[4] Several visual and cognitive factors influence the interpretation of radiographs and can enhance or hinder the accuracy of diagnosis on radiographs. [5]

In the context of radiographic diagnosis, diagnostic errors are referred to as interpretive errors.[6] Errors in interpreting diagnostic radiographs refer to mistakes made by radiologists and clinicians when evaluating radiographic images, including plain film and 3D images.[7,8] These errors can arise from flawed reasoning, incomplete knowledge, inexperience, or the clinician does not 'see' the lesion.[6,7] Interpretive errors are well documented in medical radiology, with error rates ranging from 3 to 20% globally.[7,9–15] Error rates vary depending on the type and quality of the image, the clinician's experience and expertise, cognitive biases, and mental shortcuts (heuristics). [8,9,16–18]

In dentistry, the error rate in interpreting dental radiographs is not as well documented because most radiographs are taken in a dental practice and self-reported by the dentist. [19]

Unlike medical radiologists, who undergo extensive training spanning several years, dental students and dentists typically receive less training in reading dental radiographs. This disparity in training could contribute to challenges faced by dentists in accurately interpreting images. As dental procedures become more complex, there is a greater need for precise and reliable radiographic interpretation to ensure successful treatment outcomes. [20] In busy clinical settings, where information may be ambiguous or incomplete, dentists are under pressure to make diagnostic assessments quickly, and the risk of interpretive errors increases. This pressure often stems from time constraints of standard dental appointments or patient expectations for effective resolutions of their dental issues.[21] As interpretation in radiology relies on visual perception, errors in interpretation can occur under such constrained conditions by increasing the cognitive load.[22] The multifactorial nature of these errors can lead to adverse consequences for patients and clinicians, including delayed treatment, inappropriate management plans, poor patient outcomes, and malpractice lawsuits[23,24] Understanding the role of cognitive biases is essential, as they can contribute to these interpretive errors.

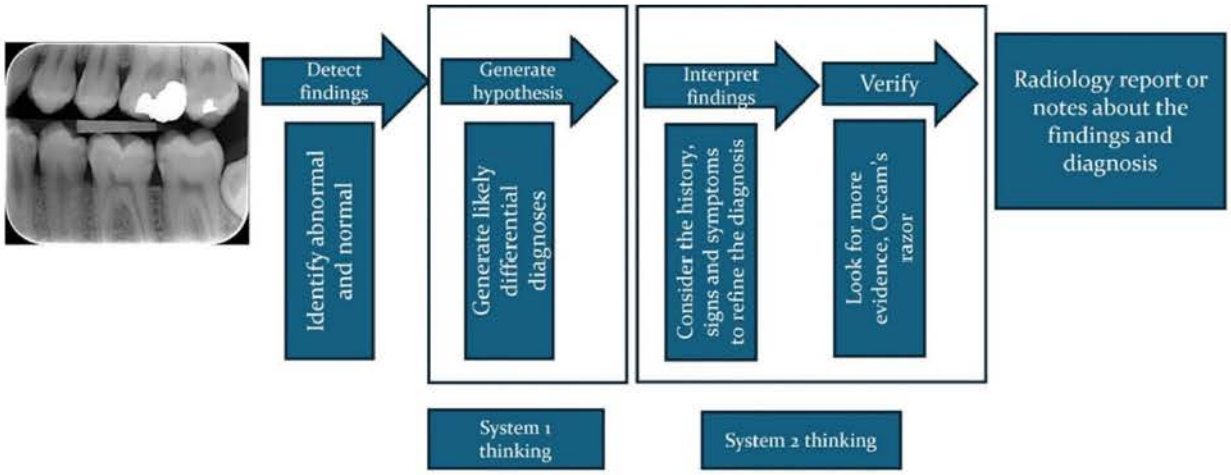


Figure 2: Diagrammatic representation of the types of thinking used during the analysis of radiographic images (Adapted from Busby LP, Courtier JL, Glastonbury CM. Bias in Radiology: The How and Why of Misses and Misinterpretations. Radiographics. 2018 Jan-Feb;38(1):236-247. doi: 10.1148/rg.2018170107. Epub 2017 Dec 1. PMID: 29194009; PMCID: PMC5790309.)

Cognitive biases are systematic patterns of deviations from the norm or rationality of judgement and result from an attempt by our brain to simplify information processing. [25,26] In medicine and dentistry, heuristic processes are generally helpful in diagnosis, but they are prone to cognitive biases. [27] These biases can lead to errors in decision-making, such as in the interpretation of dental radiographs. Kahneman and Tversky have shown that we use either system 1 or system 2 thinking in our decision-making. System 1, or fast thinking, is rapid reflexive decision-making based on heuristics, whereas system 2 thinking, also called slow thinking, involves slow, deliberate decision-making. Inexperienced clinicians may use system 2 thinking in diagnosis.[28] However, clinicians may use both system 1 and system 2 simultaneously to arrive at a diagnosis in clinical situations.[29] While these cognitive processes are helpful in diagnosis, they are susceptible to cognitive biases, which can lead to interpretive

errors.[27] Figure 2 shows how System 1 and 2 thinking may be used in the radiographic diagnostic process.

Given these challenges, there is a need for tools that can help reduce interpretive errors and improve diagnostic accuracy in dental radiology. Techniques such as artificial intelligence (AI) and machine learning (ML), reporting templates, checklists, and clinical decision-support systems have been explored.[30–35] AI and ML, in particular, have shown promise in healthcare, including dentistry and dental radiology, where all the primary data is in digital form.[36,37] Integrating AI and ML into the clinical workflow has been suggested to reduce errors in interpretations of radiographs, as these technologies can detect subtle findings on radiographs that may be missed by humans.[38] Applications of AI in imaging in healthcare are not limited to the identification of pathology on radiographs through pattern recognition but also include image enhancement to make important features on radiographs clearer and image segmentation to separate regions corresponding to normal (anatomy) and abnormal (pathology).[39–41] AI systems can mitigate cognitive biases on system 1 thinking by offering objective data-driven recommendations about the location and significance of radiographic findings. Acting as cognitive aids, they can highlight key areas on radiographs that require attention, guiding clinicians to potential pathologies or abnormalities they might otherwise overlook or miss. Thus, ML algorithms can increase efficiency and reduce errors in interpretation. Research has shown that checklists can streamline the diagnostic process and allow clinicians to focus on more critical decision-making by providing guidance and automated analysis, thereby minimising cognitive load. However, the influence of machine learning algorithms and checklists as tools to minimise the cognitive load and improve diagnostic accuracy is an unexplored area of dental research.

Gathering data on the frequency and impact of interpretive errors through a systematic review of literature and surveys[42] can provide valuable insights into the existing scientific research and dentists' understanding and perceptions of interpretive errors. This information can help us understand the role of cognitive aids, such as ML algorithms and checklists, in minimising interpretive errors and improving diagnostic accuracy in dental radiology. Further, given the potential advantages of AI technologies in dentistry, especially in dental imaging, it is crucial to understand the perceptions and attitudes of dental practitioners and dental students towards this transformative technology. By assessing these perspectives, expectations and specific concerns about AI applications can be identified and addressed, enabling stakeholders to facilitate the implementation of AI in an ethical and efficient manner.

However, the clinical implementation of AI has been limited by the 'black box' design of the algorithm, identifying a need for developing explainability methods that are easily understood and trusted by clinicians.[43] Clear explanations of ML algorithm decisions are necessary to bridge the gap between their technical complexity and their practical use in clinical settings. These insights are crucial as they will not only shape the integration of dental AI applications but also influence the development of educational programs to prepare future dentists for technologically advanced dental practice. Given these opportunities and challenges, this thesis aims to explore the role of ML algorithms as cognitive aids in reducing interpretive errors in dental radiology.

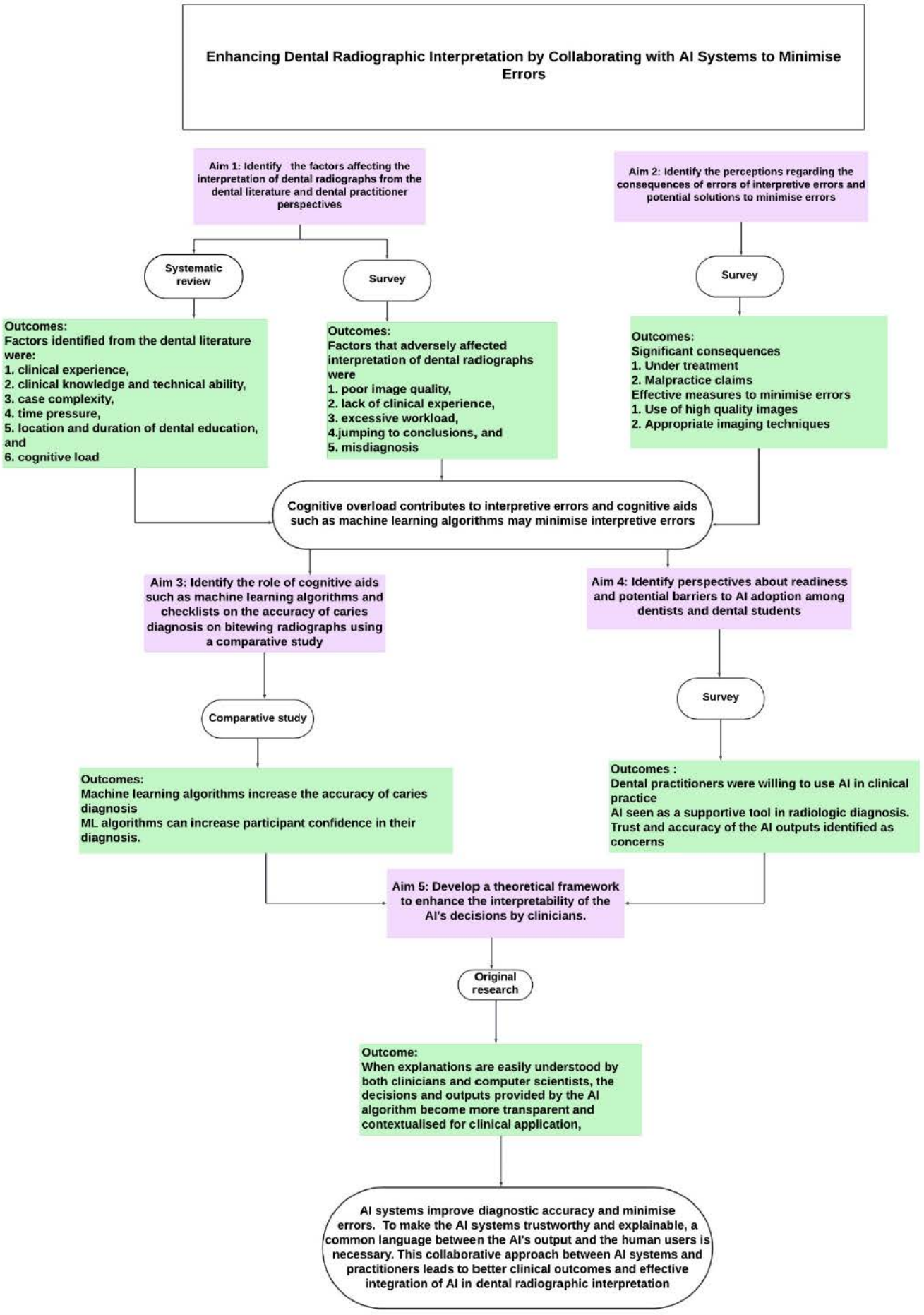

Enhancing Dental Radiographic Interpretation by Collaborating with AI Systems to Minimise Errors
Aim 1: Identify the factors affecting the interpretation of dental radiographs from the dental literature and dental practitioner perspectives
Aim 2: Identify the perceptions regarding the consequences of errors of interpretive errors and potential solutions to minimise errors
Systematic review
Survey
Survey
Outcomes:
Factors identified from the dental literature were:
1. clinical experience,
2. clinical knowledge and technical ability,
3. case complexity,
4. time pressure,
5. location and duration of dental education, and
6. cognitive load
Outcomes:
Factors that adversely affected interpretation of dental radiographs were
1. poor image quality,
2. lack of clinical experience,
3. excessive workload,
4.jumping to conclusions, and
5. misdiagnosis
Outcomes:
Significant consequences
1. Under treatment
2. Malpractice claims
Effective measures to minimise errors
1. Use of high quality images
2. Appropriate imaging techniques
Cognitive overload contributes to interpretive errors and cognitive aids such as machine learning algorithms may minimise interpretive errors
Aim 3: Identify the role of cognitive aids such as machine learning algorithms and checklists on the accuracy of caries diagnosis on bitewing radiographs using a comparative study
Aim 4: Identify perspectives about readiness and potential barriers to AI adoption among dentists and dental students
Comparative study
Survey
Outcomes:
Machine learning algorithms increase the accuracy of caries diagnosis
ML algorithms can increase participant confidence in their diagnosis.
Outcomes :
Dental practitioners were willing to use AI in clinical practice
AI seen as a supportive tool in radiologic diagnosis.
Trust and accuracy of the AI outputs identified as concerns
Aim 5: Develop a theoretical framework to enhance the interpretability of the AI's decisions by clinicians.
Original research
Outcome:
When explanations are easily understood by both clinicians and computer scientists, the decisions and outputs provided by the AI algorithm become more transparent and contextualised for clinical application,
AI systems improve diagnostic accuracy and minimise errors. To make the AI systems trustworthy and explainable, a common language between the AI's output and the human users is necessary. This collaborative approach between AI systems and practitioners leads to better clinical outcomes and effective integration of AI in dental radiographic interpretation

Figure 1: Concept map depicting the trajectory of the thesis

1.3 Research objectives

Research question	Chapters
Research question 1: What are the factors affecting the interpretation of dental radiographs?	2,3
Research question 2: What are the consequences of interpretive errors and potential measures to minimise them?	4
Research question 3: What is the role of cognitive aids such as machine learning (ML) algorithms and checklists in improving diagnostic accuracy and minimising interpretive errors?	5
Research question 4: What are the attitudes and perceptions of Australian dentists and students towards dental AI applications?	6
Research question 5: How can the outputs of ML algorithms be made easily understood and interpretable by end users such as dentists?	7

Table 1: Research questions and the chapters where they are addressed

The primary objective of this thesis was to investigate whether augmenting dentists with artificial intelligence (AI) systems can enhance the interpretation of dental radiographs by improving diagnostic accuracy. The research questions and their relationship to the articles are presented in Table 1.

To achieve this objective, the research explored five interrelated areas to provide a better understanding of the potential of AI to improve dental radiographic interpretation. The two central themes that guided the research were: 1. The various factors that influence the interpretation of dental radiographs, the occurrence of interpretive errors and the consequences of such errors. 2. The role of cognitive aids such as ML algorithms and diagnostic checklists in minimising interpretive errors and improving diagnostic accuracy. Additionally, this research explored the attitudes and perceptions of Australian dentists and dental students towards adopting AI in dental practice. This work also proposed a new explainability method to address the challenge of making the outputs of ML algorithms more accessible and interpretable to end users, such as dentists.

The initial focus was identifying the factors affecting the interpretation of dental radiographs, which was accomplished through a review of existing literature and a survey of dental practitioners to gather their perceptions. A systematic review of dental literature was conducted to determine what is known about these factors. A survey was conducted among dental practitioners in New South Wales (NSW) to assess their knowledge and perceptions regarding factors affecting interpretive errors of dental radiographs. The systematic review of the literature is presented in **Chapter 2**, and the survey is presented in **Chapter 3**.

A general understanding of the factors that influence the interpretation of dental radiographs was essential, as this formed the foundation for identifying the consequences and potential measures to minimise interpretive errors. These aspects were addressed through a survey of dental practitioners, which explored their perceptions about the consequences of interpretive errors for clinicians and patients and possible measures to minimise them. The survey is presented in **Chapter 4**.

The role of cognitive aids, such as ML algorithms and checklists, in reducing errors and improving diagnostic accuracy was assessed. This part of the study examined the comparative effectiveness of ML and diagnostic checklists in enhancing the accuracy of caries diagnosis on bitewing radiographs. This study is presented in **Chapter 5**.

Understanding the attitudes of dental students and practitioners towards AI applications in dentistry is essential, as the ease of integrating AI tools into clinical practice is likely influenced by their acceptance and willingness to use such tools. To gain insight into these perspectives, a survey was conducted among dental students and practitioners in Australia, focusing on their perceptions about the benefits and concerns associated with dental AI applications. The findings of this study are presented in **Chapter 6**.

The survey of dental students and dental practitioners regarding applications of AI in dentistry recognised that for AI tools to be practically useful, their outputs must be accessible and interpretable by the end users. To address this need, an original research study was conducted to explore how the outputs of ML algorithms can be made more accessible and interpretable for end users. This study proposes a theoretical explainability framework that clinicians and other healthcare end users of ML algorithms can understand. This original research led to the development of a novel explainability framework grounded in image features such as density and distance. This framework was designed to be easily understood by both clinicians who use AI in clinical practice and the developers who create the AI algorithms. The study is presented in **Chapter 7**.

1.4 Grants:

This work was partially supported by a grant from the Australian Dental Research Foundation (grant number 2790/2020).

1.5 References:

1. White SC, Heslop EW, Hollender LG, Mosier KM, Ruprecht A, Shrout MK. Parameters of radiologic care: An official report of the American Academy of Oral and Maxillofacial Radiology. Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology [Internet]. 2001;91(5):498–511. Available from: https://www.sciencedirect.com/science/article/pii/S1079210401271327

2. White SC, Yoon DC, Tetradis S. Digital Radiography in Dentistry: What It Should Do for You. J Calif Dent Assoc [Internet]. 1999 Dec 1;27(12):942–52. Available from:

https://doi.org/10.1080/19424396.2016.12221191

3. Vandenberghe B, Jacobs R, Bosmans H. Modern dental imaging: a review of the current technology and clinical applications in dental practice. Eur Radiol [Internet]. 2010 Nov;20(11):2637–55. Available from: https://www.proquest.com/scholarly-journals/modern-dental-imaging-review-current-technology/docview/756902449/se-2?accountid=14757

4. Whaites E, Drage N. Essentials of Dental Radiography and Radiology. London: Elsevier Health Sciences; 2013.

5. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. Am J Roentgenol. 2013;201(3):611–7.

6. Onder O, Yarasir Y, Azizova A, Durhan G, Onur MR, Ariyurek OM. Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review. Insights Imaging [Internet]. 2021;12(1):51. Available from: https://doi.org/10.1186/s13244-021-00986-8

7. Bruno MA, Walker EA, Abujudeh HH. Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction. Radiographics [Internet]. 2015;35(6):1668–76. Available from: http://pubs.rsna.org/doi/10.1148/rg.2015150023

8. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive Error in Radiology. AJR Am J Roentgenol. 2017 Apr;208(4):739–49.

9. Kim, YW ML. Fool me Twice: Delayed Diagnosis in radiology. Am J Roentgenol. 2014;202:465–70.

10. Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. World J Radiol [Internet]. 2010 Oct;2(10):377–83. Available from: http://www.wjgnet.com/1949-8470/full/v2/i10/377.htm

11. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol. 1978;13(3):175–81.

12. Berlin L. Accuracy of Diagnostic Procedures: Has It Improved Over the Past Five Decades? Am J Roentgenol [Internet]. 2007 May 1;188(5):1173–8. Available from: https://doi.org/10.2214/AJR.06.1270

13. Garland HL. On the scientific evaluation of diagnostic procedures. Radiology. 1949;52(3):309–28.

14. Graber ML, Kissam S, Payne VL, Meyer AND, Sorensen A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: A narrative review. BMJ Qual Saf. 2012;21(7):535–57.

15. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. BMJ Qual Saf. 2017;26(6):484–94.

16. Krupinski EA. Current perspectives in medical image perception. Atten Percept Psychophys. 2010 Jul;72(5):1205–17.

17. Iskander M. Burnout, Cognitive Overload, and Metacognition in Medicine. Med Sci Educ [Internet]. 2019;29(1):325–8. Available from: https://doi.org/10.1007/s40670-018-00654-5

18. Weber LS, Schulze R. The effects of examiner fatigue on the diagnostic accuracy of dental radiographs. Clin Oral Investig. 2021 Nov;25(11):6193–9.

19. Choi J. Dental radiology reporting status and recording frequency of reporting items in Korea. Imaging Sci Dent. 2023 Mar;53(1):35–42.

20. Stefanac SJ, Nesbit SP. Diagnosis and Treatment Planning in Dentistry: Diagnosis and Treatment Planning in Dentistry - E-Book [Internet]. 4th editio. Nesbit S, Stefanac S, editors. Elsevier Health Sciences; 2023. 480 p. Available from: https://books.google.com.au/books?id=HcCrEAAAQBAJ

21. Plessas A, Nasser M, Hanoch Y, O'Brien T, Bernardes Delgado M, Moles D. Impact of time pressure on dentists' diagnostic performance. J Dent [Internet]. 2019;82(January):38–44. Available from: https://doi.org/10.1016/j.jdent.2019.01.011

22. Waite S, Grigorian A, Alexander RG, Macknik SL, Carrasco M, Heeger DJ, et al. Analysis of Perceptual Expertise in Radiology – Current Knowledge and a New Perspective. Front Hum Neurosci [Internet]. 2019;13. Available from: https://www.frontiersin.org/article/10.3389/fnhum.2019.00213

23. Nikdel C, Nikdel K, Ibarra-Noriega A, Kalenderian E, Walji MFMF. Clinical dental faculty members' perceptions of diagnostic errors and how to avoid them. J Dent Educ [Internet]. 2018 Apr 1;82(4):340–8. Available from: http://www.jdentaled.org/content/82/4/340.abstract

24. White BA, Maupomé G. Making clinical decisions for dental care: concepts to consider. Spec Care Dent [Internet]. 2003 Sep 1;23(5):168–72. Available from: https://doi.org/10.1111/j.1754-4505.2003.tb00306.x

25. Yoon SY, Lee KS, Bezuidenhout AF, Kruskal JB. Spectrum of Cognitive Biases in Diagnostic Radiology. RadioGraphics [Internet]. 2024 Jun 6;44(7):e230059. Available from: https://doi.org/10.1148/rg.230059

26. Tee QX, Nambiar M, Stuckey S. Error and cognitive bias in diagnostic radiology. J Med Imaging Radiat Oncol [Internet]. 2022 Mar 1;66(2):202–7. Available from: https://doi.org/10.1111/1754-9485.13320

27. Busby LP, Courtier JL, Glastonbury CM. Bias in Radiology: The How and Why of Misses and Misinterpretations. Radiogr a Rev Publ Radiol Soc North Am Inc. 2018;38(1):236–47.

28. Kahneman D. Thinking, fast and slow. Macmillan; 2011.

29. Tay SW, Ryan P, Ryan CA. Systems 1 and 2 thinking processes and cognitive reflection testing in medical students. Can Med Educ J. 2016 Oct;7(2):e97–103.

30. Wang CW, Huang CT, Lee JH, Li CH, Chang SW, Siao MJ, et al. A benchmark for comparison of dental radiography analysis algorithms. Med Image Anal. 2016;31:63–76.

31. Krombach JW, Edwards WA, Marks JD, Radke OC. Checklists and Other Cognitive Aids For Emergency And Routine Anesthesia Care-A Survey on the Perception of Anesthesia Providers From a Large Academic US Institution. Anesthesiol pain Med. 2015 Aug;5(4):e26300.

32. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. Acad Med. 2011;86(3):307–13.

33. Winters BD, Aswani MS, Pronovost PJ. Commentary: Reducing diagnostic errors: Another role for checklists? Acad Med. 2011;86(3):279–81.

34. Wenzel A. Computer-Automated Caries Detection in Digital Bitewings: Consistency of a Program and Its Influence on Observer Agreement. Caries Res. 2001;35(1):12–20.

35. de Bruijne M. Machine learning approaches in medical image analysis: From detection to diagnosis. Medical Image Analysis. 2016;33.

36. Fatima A, Shafi I, Afzal H, Díez IDLT, Lourdes DRSM, Breñosa J, et al. Advancements in Dentistry with Artificial Intelligence: Current Clinical Applications and Future Perspectives. Healthc (Basel, Switzerland). 2022 Oct;10(11).

37. Heo MS, Kim JE, Hwang JJ, Han SS, Kim JS, Yi WJ, et al. Artificial intelligence in oral and maxillofacial radiology: what is currently possible? Dentomaxillofac Radiol. 2021 Mar;50(3):20200375.

38. Najjar R. Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. Diagnostics (Basel, Switzerland). 2023 Aug;13(17).

39. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer [Internet]. 2018;18(8):500–10. Available from: http://dx.doi.org/10.1038/s41568-018-0016-5

40. Valavanidis A. Artificial Intelligence in Medical Diagnostics and Imaging. Applications that will revolutionise the fields in biomedical research and healthcare. 2023 Nov 17;1:1–38.

41. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Vol. 6, Future healthcare journal. 2019. p. 94–8.

42. Balogh EP, Miller BT, Ball JR, The National Academies of Sciences, Engineering and M. Improving Diagnosis in Health Care [Internet]. Balogh EP, Miller BT, Ball JR, editors. Improving Diagnosis in Health Care. Washington, DC: The National Academies Press; 2015. 1–472 p. Available from: https://www.nap.edu/catalog/21794/improving-diagnosis-in-health-care

43. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. Cureus. 2023 Oct;15(10):e46454.

Chapter 2

Factors Affecting the Interpretation of Dental Radiographs - Systematic Review

This systematic review identifies and analyses, from the dental literature, the factors contributing to errors in interpreting dental radiographs. Understanding these factors is crucial for developing strategies to minimise errors and enhance the accuracy of radiographic interpretations in dental practice.

Dentomaxillofacial Radiology (2022) 51, 20220279
© 2022 The Authors. Published by the British Institute of Radiology under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License http://creativecommons.org/licenses/by-nc/4.0/, which permits unrestricted non-commercial reuse, provided the original author and source are credited.

birpublications.org/dmfr

SYSTEMATIC REVIEW

Factors affecting interpretation of dental radiographs

[1]Shwetha Hegde, [2]Jinlong Gao, [3]Rajesh Vasa and [4]Stephen Cox

[1]Academic Fellow, Dentomaxillofacial Radiology, Sydney Dental School, University of Sydney, Sydney, Australia; [2]Senior Lecturer, Sydney Dental School, Institute of Dental Research, Westmead Centre for Oral Health, University of Sydney, Sydney, Australia; [3]Head of Translational Research and Development, Applied Artificial Intelligence, Deakin University, Melbourne, Australia; [4]Head of Discipline, Discipline of Oral Surgery, Sydney Dental School, University of Sydney, Sydney, Australia

Objectives: To identify the factors influencing errors in the interpretation of dental radiographs.
Methods: A protocol was registered on Prospero. All studies published until May 2022 were included in this review. The search of the electronic databases spanned Ovid Medline, PubMed, EMBASE, Web of Science and Scopus. The quality of the studies was assessed using the MMAT tool. Due to the heterogeneity of the included studies, a meta-analysis was not conducted.
Results: The search yielded 858 articles, of which eight papers met the inclusion and exclusion criteria and were included in the systematic review. These studies assessed the factors influencing the accuracy of the interpretation of dental radiographs. Six factors were identified as being significant that affected the occurrence of interpretation errors. These include clinical experience, clinical knowledge, and technical ability, case complexity, time pressure, location and duration of dental education and training and cognitive load.
Conclusions: The occurrence of interpretation errors has not been widely investigated in dentistry. The factors identified in this review are interlinked. Further studies are needed to better understand the extent of the occurrence of interpretive errors and their impact on the practice of dentistry.
Dentomaxillofacial Radiology (2022) **51**, 20220279. doi: 10.1259/dmfr.20220279

Cite this article as: Hegde S, Gao J, Vasa R, Cox S. Factors affecting interpretation of dental radiographs. *Dentomaxillofac Radiol* (2022) 10.1259/dmfr.20220279.

Keywords: errors; interpretation; dental radiology

Introduction

Making a diagnosis is a critical clinical decision and has implications for clinicians and their patients. The diagnostic process involves gathering data from patient history and clinical examination, performing diagnostic tests, and arriving at a diagnosis by interpreting and integrating the findings.[1]

In clinical dentistry, radiographs provide useful diagnostic information about diseases and pathologies of the teeth and jaws. They are routinely used in the clinical diagnosis of problems and contribute to treatment planning.[2,3] Evaluation of a radiographic image involves visual inspection and interpretation.[4] Making a diagnosis from an image requires four steps: (a) *detection*- identifying a finding that would require further analysis. (b) *recognition*-identification of pathology, (c) *discrimination*- characterisation of the lesion, and (d) *diagnosis and differential diagnosis*.[5] The detection and discrimination of pathology in a radiographic image involves both perceptual (recognising a difference or change in appearance) and cognitive (understanding the significance of these changes) processes. The clinician must first detect pathology on a radiograph and then characterise it, leading to a diagnosis. However, this is a complex process, and decisions are often made based on incomplete clinical information. In some situations, inaccurate thinking processes can lead to errors in diagnosis. Due to the analytical complexity of working with partial information, diagnostic errors and errors of

Correspondence to: Dr Shwetha Hegde, E-mail: shwetha.hegde@sydney.edu.au; drsshegde@gmail.com

Received 24 August 2022; revised 24 November 2022; accepted 25 November 2022

interpretation of radiographs are frequently unavoidable in day-to-day practice.[6]

Errors of interpretation have been described as a discrepancy in the interpretation that significantly differs from the consensus of one's peers.[7] The errors in interpreting radiographs have been studied in medicine.[8–12] While interpretive errors and the factors affecting them could be the same in dentistry, they have not been assessed with the same rigour as in medicine and are therefore not well understood.

Aims

This systematic review aimed to identify the factors contributing to errors in interpreting radiographs by dental practitioners.

Objectives

(1) Identify factors that influence or cause the errors in interpretation
(2) Identify the relative significance of those factors

Methods

Protocol and registration

This systematic review followed PRISMA guidelines[13] and was registered on the National Institute of Health Research Database (www.crd.york.ac.uk/prospero, protocol registration number CRD42020207998).

Eligibility criteria

Inclusion criteria

(1) Studies that examined factors influencing interpretation errors or diagnostic errors among dental professionals
(2) Published in the English language
(3) Articles published in peer-reviewed journals.
(4) All clinical settings.
(5) Reference lists of selected papers from the search

Exclusion criteria:

1. Case reports, conference proceedings, letters to the editor and news articles

Data sources

The search focused on the factors influencing interpretation errors of radiographs in clinical environments with no restrictions to study design. Ovid Medline, PubMed, EMBASE, PsycINFO, the Cochrane Library (Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane methodology register), Scopus, Web of Science (including Science Citation Index Expanded) and EBSCO databases were systematically searched. The search was conducted between February 2021 and March 2021 and then updated in May 2022. The databases were searched for articles published between 1946 and May 2022. The search was limited to articles published in English.

Search strategy

The search strategy was developed to include relevant concepts specific to dental radiographic interpretation errors. The databases were searched using keywords, Medical Subject Headings (MeSH), Emtree terms and free-text terms relating to radiology, dentomaxillofacial radiology, dentistry, errors of interpretation, diagnostic errors and clinical decision making. The Ovid Medline search used a combination of Mesh terms, and the same strategy was used for searching other databases. The Ovid database search strategy is included in Table 1. The potential factors affecting errors of interpretation were identified from the medical radiology literature and adapted to the dental context. The search strategy was validated by checking if seminal papers were captured in the database search. The conceptual structure of the search strategy is described in Table 2.

Data collection process and data items

Full texts of those articles that met the inclusion criteria were obtained, and data were extracted. A data extraction template was developed on Covidence (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org). The data extraction form was tested using two included papers, and necessary modifications were made before extraction. Two reviewers (SH and JG) independently extracted the data using Covidence, and any conflict was resolved through discussion.

Table 1 search strategy

Dentistry/ or dentists/ or dentist or medical radiologist* or dental radiolog* or dentomaxillofacial radiolog* or dental imaging or dental X-ray* or radiographer* or panoramic radiograph* or OPG or orthopantomograph* or intraoral radiograph* or intraoral radiograph*).mp.*
Anxiety/ or Burnout, Professional/ or Occupational Stress/px or fatigue or anxiety or depression or stress or burnout or workload or occupational stress or occupational burnout or job-related stress or job stress* or human error* or professional error* or time pressure or "not taking adequate time" or professional experience.mp.
Clinical Decision-Making/ or Decision Making/ or Diagnostic Self Evaluation/ or Diagnostic Errors/ or (diagnostic error* or misdiagnos#s or mis-diagnos#s or missed diagnos#s or delayed diagnosis or diagnos* mistake* or radilogical error* or interpretive error*).mp.

Table 2 Conceptual structure of search strategy

Concept 1	*Concept 2*	*Concept 3*
dentist and dental imaging	factors causing errors of interpretation	consequences
Keywords: dentist, dentomaxillofacial radiology, dental radiology, dental imaging, panoramic, intraoral imaging	Keywords: stress, occupational stress, fatigue, burnout, excessive workload, attentional bias, case complexity, inattention, time pressure, clinical experience	Keywords: errors in decision making, problems in diagnosis, diagnostic errors, missed diagnosis, errors of omission, interpretive errors, errors of interpretation clinical decision making, image perception, delayed diagnosis, observer variation, radiological error,
MeSH terms: "Dentists"[Mesh] OR "Radiography, Dental, Digital"[Mesh] OR "Radiography, Panoramic"[Mesh]	MeSH terms: "Occupational Stress"[Mesh] OR "Stress, Psychological"[Mesh] OR "Fatigue"[Mesh] OR "Burnout, Professional"[Mesh] OR "Workload"[Mesh] OR "Attentional Bias"[Mesh]	MeSH terms: ((((("Diagnostic Errors"[Mesh]) OR "Missed Diagnosis"[Mesh]) OR "Medical Errors"[Majr]) OR "Medical Errors"[Mesh]) OR "Clinical Decision-Making"[Mesh]) OR "Delayed Diagnosis"[Mesh]) OR "Observer Variation"[Mesh]

Methodological quality

The methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT).[14] The MMAT is a validated critical appraisal tool designed to appraise systematic mixed studies reviews. It permits the assessment of the methodological quality of five categories of studies: qualitative research, randomised controlled trials, non-randomised studies, quantitative descriptive studies, and mixed methods studies.[14] One reviewer (SH) independently performed the quality assessment, and then the second reviewer (JG) validated the information, and any disagreement was resolved through discussion.

Depending on the MMAT criteria, papers were grouped into high, medium, or low quality. For each study, the overall quality score was calculated using the MMAT. Based on the MMAT tool, studies were scored as unclassified, 20%, 40%, 60%, 80%% or 100%, with scores of 80–100% considered high quality.[14]

Publication and reporting bias were minimised by searching multiple databases and including studies from those sites. Studies with low methodological quality were not excluded but were described in the analysis and synthesis.

Results

The results of this review were tabulated and described narratively. The eight studies in the systematic review highlighted the following factors as being significant for interpretation errors of dental radiographs: clinical experience, clinical knowledge and technical ability, time pressure, cognitive load, case complexity on dental radiographs, and training (duration of the didactic program). One study[15] did not fit these categories and was described separately.

Meta-analyses were not attempted due to the heterogeneity of the raw data due to widely varying study settings and study designs.

Study selection

Eight hundred fifty-eight articles were obtained from searching the databases and reference lists of selected papers. These articles were imported into Covidence for screening and data extraction. After the duplicates were removed, 336 articles were available for the title and abstract screening. Two reviewers (SH and JG) independently screened the titles and abstracts to identify studies that met the inclusion criteria. All studies deemed eligible by both reviewers were subjected to a full article review to determine inclusion in the final analysis, and eight articles met the inclusion criteria. They were included in the final data extraction. Any differences about the eligibility of inclusion were resolved by discussion between the two reviewers. The screening process is shown in the PRISMA flow diagram,[13] Figure 1.

Study characteristics

An overview of the eight studies included in this review is shown in Table 3.

Participants and settings

Dentists (clinicians) and dental students were participants in five of the included studies.[16–20] One study was a retrospective analysis of dental records.[15] The study location varied among the included articles, with two studies based in the USA,[16,17] two in Brazil,[18,20] one in the Netherlands,[16] one in Germany,[21] one in the UK,[19] and 1 in two locations-Brazil and Switzerland.[22] 5 of the studies used radiographs of extracted teeth.[16–18,20,22] Two studies used patient radiographs.[19,22] The outcomes were similar for five studies, and they examined the diagnostic accuracy of dental caries on bitewing radiographs. One study[19] used panoramic radiographs for patients, and the other[21] used CBCT images to detect MB2 of maxillary molars. The final study[15] used malpractice claims to assess the extent of preventable adverse events and included radiographic interpretation errors.

The majority (62%) of the studies were conducted entirely in a university clinic setting,[17–21] with one using both private practice and a university setting.[16] One study[15] was a retrospective audit conducted by a panel of dentists and dentists and physicians with legal training in private practice.

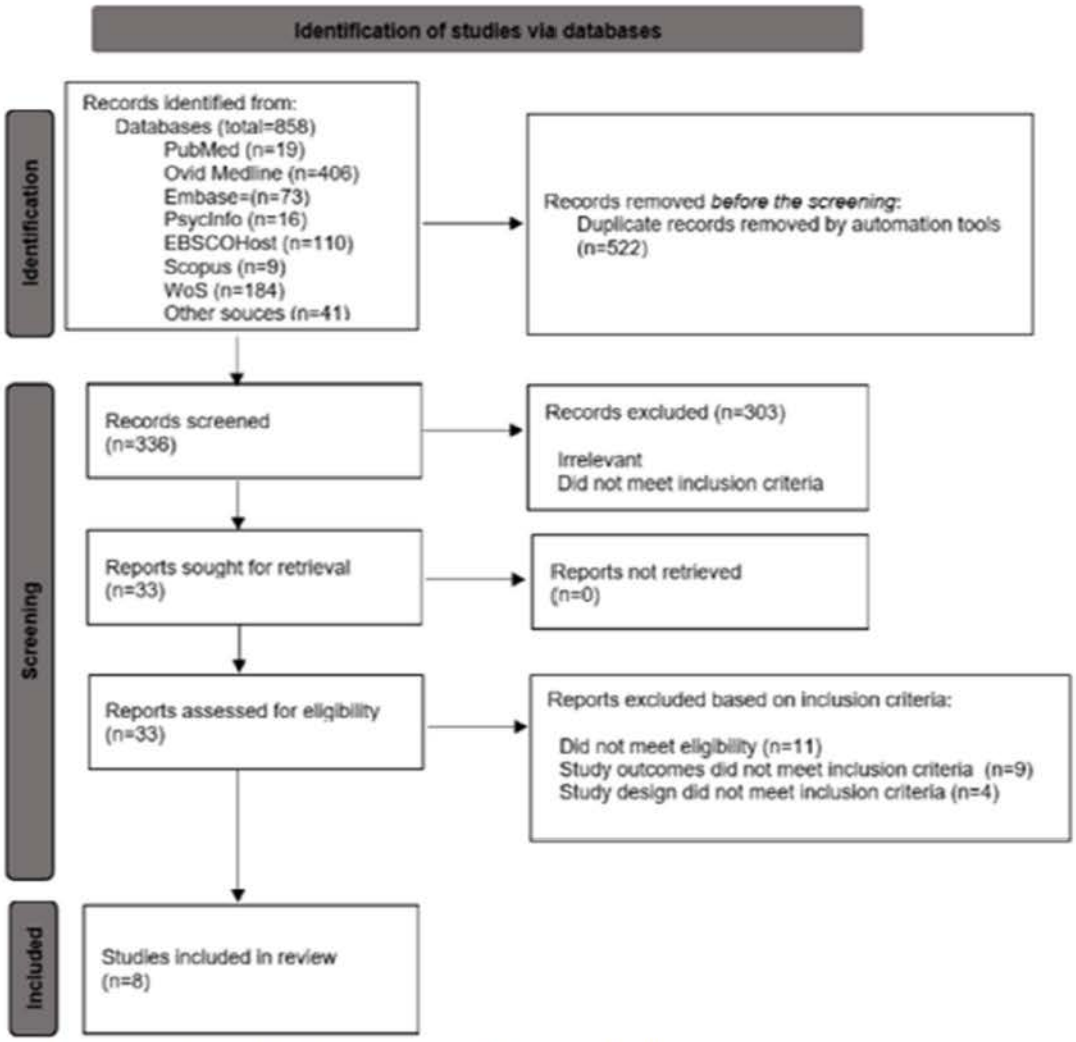


Figure 1 Prisma flow diagram for study identification, screening, eligibility, and inclusion.

Quality assessment

The quality assessment of all included studies is shown in Table 4. This review has three categories of studies: non-randomised quantitative studies,[16–19,21], a quantitative randomised cross-over study[22] and a quantitative descriptive study.[15] In the MMAT checklist, there were five questions for each study type, to which the response was 'yes', 'no' or 'can't tell'.

Seven[15–20,22] of the eight studies were deemed medium quality with a score of 60%, and one study[21] was rated low quality with a score of 40%.

The participants did not represent the target population in four studies due to limited sample sizes.[17–19,21] While all seven studies[16–22] had dentists and dental students as participants, the study design did not specify the inclusion and exclusion criteria used, what methods were used to achieve a sample size of participants representing the target population, sample size calculation and whether the sample size was adequate for the target population. In addition, no attempt was made to minimise confounding bias either by using large sample sizes, stratification, or statistical control.

Five of the six non-randomised quantitative studies[16–19,21] conducted measurements for outcomes and intervention (or exposure). Clinical experience was considered as exposure in these studies. The outcome was the diagnostic accuracy of radiographic findings and was measured using sensitivity and specificity of diagnosis. However, none of the five studies in this category accounted for confounders in design and analysis, such as factors influencing clinical experience, namely type of clinical practice, the dental clinic's location (urban, rural), number of patients seen each day, number of radiographs diagnosed each day. The influence of changing diagnostic standards and criteria and evolving scientific knowledge on clinical experience were also not considered.

The quantitative randomised cross-over study by Plessas et al. (2019)[22] did not include blinding in the study design, the outcome assessors (participants) were not

Table 3 data extraction table

	Author, year and country	*Aim*	*Study design*	*Participant sample size*	*Image type and image sample size*	*Participants*	*Exposure (factor investigated)*	*Comparator*	*Outcome (Effects of the factor studied)*	*Conclusion*
1	[17](17), The Netherlands	To compare the diagnostic accuracy of dental students and general dental practitioners	Nonrandomised quantitative	Dentists = 273 Dental students = 259	Bitewing radiographs of extracted teeth showing 105 interproximal surfaces	Dentists	Clinical experience Years of experience = not stated	fourth-year dental students from three consecutive cohorts	diagnostic accuracy of dentinal caries on bitewing radiographs	Clinical experience increased specificity but inversely affected the sensitivity of radiographic diagnosis of dental caries, with dental students having higher sensitivity scores than dentists.
2	[18](18), Brazil and Switzerland	To assess the influence of clinician experience on the reproducibility and accuracy of radiographic examination for detection of occlusal caries.	Nonrandomised quantitative study	Dentists = 10 Dental students = 10	Bitewing radiographs of 166 extracted permanent teeth	Brazilian and Swiss dentists	clinical experience Years of experience of dentists = 5 to 7 years location of dental education/ training programs	Brazilian and Swiss dental students	accuracy (reproducibility and validity) of caries diagnosis at two thresholds	The clinical experience affected the sensitivity and specificity of radiographic caries diagnosis Clinical experience increased specificity. This paper did not specify how the education program (location) affected diagnostic accuracy
3	[16](16), USA	To identify and distinguish treatment related injuries that are avoidable from those that are sequelae of a dental problem	Retrospective cohort study (qualitative)	92 dental malpractice claims Poor interpretation and incorrect prescription of radiographs = 8 cases	not applicable	92 dental malpractice claims filed between 1900 and 1974	Knowledge and technical ability in radiography	No comparator	Dental adverse events and degree of avoidability of adverse events	Poor quality films and interpretation resulting in diagnostic errors were considered avoidable outcomes.

(Continued)

Table 3 (Continued)

	Author, year and country	*Aim*	*Study design*	*Participant sample size*	*Image type and image sample size*	*Participants*	*Exposure (factor investigated)*	*Comparator*	*Outcome (Effects of the factor studied)*	*Conclusion*
4	[19](19), USA	To compare the effect of differing levels of didactic education and clinical experience among first- and senior-year dental students and the dental faculty on the diagnosis of occlusal caries on radiographs	Nonrandomised quantitative	Dentists = 15 Dental students in the first year = 15 Dental students in senior year = 14	Bitewings of 100 extracted teeth for assessment of occlusal caries	Dental faculty	clinical experience Years of experience of the dental faculty-17 years levels of didactic education (first and senior year students)	first-year and senior-year dental students	Confidence of presence/absence of caries using a 5-point Likert scale)	Clinical experience affected the sensitivity and specificity of caries diagnosis, with the specificity of radiographic caries diagnosis increasing with increasing experience. There was no significant difference in levels of accuracy between the first year and senior students
5	[20](20), Germany	To compare the pupillary responses (pupil diameter, gaze hit mapping) of experts and student dentists to panoramic radiographs of varying difficulties.	Nonrandomised quantitative study	Dentists = 26 Dental students = 50	20 patient panoramic radiographs for students and 15 for experts	Dentists (GP and specialists) working at the university clinics	Cognitive load clinical experience case (image) difficulty/ complexity years of experience of dentists = 10 years on average	Third-year dental students	Cognitive load was measured using pupil diameter. An increase in median pupillary diameter from baseline and variation with the level of difficulty on the radiograph was measured	More experienced clinicians show a proportional increase in pupillary response to increasing case difficulty. Dental students had a inconsistent increase in pupillary diameter regardless of difficulty level.
6	[23](21), Brazil	To assess the ability of three observers to accurately confirm the existence and absence of MB2 canals in human first upper molars with different root conditions and the agreement among three observers.	Nonrandomised quantitative study	Dentists (OMR registrars) =3	CBCT of 82 extracted maxillary first molars was used. The teeth were grouped into three depending on the condition of MB1 as non-filled, filled and deobturated	Dentists (OMR registrars) in the oral and maxillofacial radiology program.	clinician experience Years of experience = reviewer one five years' experience, reviewer two three years' and reviewer three one year	Comparison of interobserver agreement	Agreement on identifying the presence of MB2 (second mesiobuccal canal) under different conditions of MB1 (non-filled, filled, and deobturated)	Agreement of MB2 diagnosis increased with clinical experience

(Continued)

Table 3 (Continued)

	Author, year and country	*Aim*	*Study design*	*Participant sample size*	*Image type and image sample size*	*Participants*	*Exposure (factor investigated)*	*Comparator*	*Outcome (Effects of the factor studied)*	*Conclusion*
7	[22](22), UK	To explore the role of time pressure on the radiographic diagnostic performance when viewing bitewing radiographs among dentists	Randomised cross-over study	Dentists = 40 randomly assigned to four groups of ten each	12 patient bitewing radiographs Each bitewing had a range of difficulty (three easy, three difficult)	Primary care dentists (GP) Years of experience = 17 years (average)	time pressure	time pressure versus no time pressure time	radiographic diagnostic performance-diagnostic error	Under time pressure, the sensitivity was lower, but the specificity was not affected. Time pressure increased the incidence of diagnostic errors and errors of omission.
8	Bussaneli DG et al, 2014 (23), Brazil	To evaluate the influence of the examiner's clinical experience on the detection and treatment decision of caries lesions in primary molars.	Nonrandomised quantitative study	Dental faculty = 3 Dental students = 3	Bitewing radiographs of 77 recently extracted or exfoliated primary molars.	Dentists from the department of paediatric dentistry.	clinical knowledge (ICDAS classification on radiographs) Clinical experience Years of experience of the dental faculty = 10 years (group A). Dental students (group B) were familiar with the ICDAS criteria	Second-year dental students	Accuracy of caries diagnosis using clinical evaluation and bitewing radiographs.	The professional experience did not affect the accuracy of caries diagnosis on bitewings.

Table 4 Quality assessment

	Quantitative randomised Cross-over					*Quantitative non-randomised*					*Quantitative descriptive*					*Quality*
References	2.1	2.2	2.3	2.4	2.5	3.1	3.2	3.3	3.4	3.5	4.1	4.2	4.3	4.4	4.5	
Mileman 2002						*	*	*	-	-						medium
Diniz 2010						-	*	*	-	*						medium
Lazarchik 1995						-	*	*	-	*						medium
Castner 2020						-	*	*	-	*						medium
Vizzotto 2015						-	-	*	-	*						low
Bussaneli 2015						-	*	*	Can't tell	*						medium
Plessas 2019	*	*	*	-	Can't tell											medium
Milgrom 1975											*	*	Can't tell	*	-	medium

1*meets 20% of MMAT criteria
2**meets 40% of MMAT criteria
3***meets 60% of MMAT criteria
4****meets 80% of MMAT criteria
5*****meets 100%of MMAT criteria
20-40%=low quality; 60–80% = medium quality; 80–100% = high quality

blinded to the intervention, and it was unclear whether the participants adhered to the assigned intervention.

The quantitative descriptive study[15] did not describe a sampling strategy and did not contain statistical analysis, and it was unclear whether the outcome measurements were appropriate. The variables were defined, but there was no description of how the variables were measured. The quality assessment of all studies using MMAT[14] is shown in Table 4.

Results of Synthesis

This review identified six factors influencing interpretation errors: the complexity of dental radiographs, clinical knowledge and technical ability, clinical experience, cognitive load, time pressure and dental training and education programs. These factors appear interrelated and contribute to the occurrence of interpretation errors.

The complexity of dental radiographs

The complexity of a radiograph describes the level of difficulty in detecting pathology and interpreting the findings. The complexity of a radiograph depends on the type (bitewings, panoramic radiographs, CBCT) and quality of radiographs (image contrast, presence of film faults), the type and incidence of pathology identified on the radiographs, and the conspicuity of the findings.[23] This review found that interpretation errors were more likely to occur with radiographs of higher complexity.

In the studies in this systematic review, dental radiographs with various complexities were considered to evaluate the influence of complexity on interpretation errors. Most studies (50 %) used *in vitro* bitewing radiographs of extracted teeth to evaluate a single type of pathology, dental caries.[16–18,20] The study by Plessas et al[22] used patient bitewing radiographs, and pathological findings included dental caries and periodontal disease.

In contrast, Vizzotto et al[21] utilised cone beam computed tomography (CBCT) images of extracted maxillary molars to study the diagnostic accuracy of identifying the second mesiobuccal canal (MB2) of maxillary molars. The study by Castner et al[19] directly assessed the effect of case complexity on diagnostic accuracy. They used patient panoramic radiographs and categorised the images depending on the difficulty of interpretation, which relied on the prevalence of the lesions and ease of detection (lesion conspicuity). They found that the diagnostic accuracy depended on the level of complexity of findings on the radiographs.

The seven studies described here measured the diagnostic accuracy using sensitivity and specificity. The type of image- bitewings, panoramic radiographs or CBCT, *in vitro* or patient radiographs and a variety of pathologies studied varied in the studies included in this review. These factors can affect the case complexity of dental radiographs and influence the accuracy of diagnosis and the occurrence of interpretation errors. In a clinical situation, the interpretation of patient radiographs may have a higher complexity than the radiographs of extracted teeth (*in vitro* images) used in experimental settings, as patient radiographs may show both expected and incidental findings. In addition to the artificial reading environment, this could influence diagnostic accuracy.

Clinical knowledge and technical ability in dental radiography

Clinical knowledge is described as scientific knowledge about diseases, their pathophysiology, bodily processes, appropriate diagnostic tests and therapeutic measures.[24] In contrast, technical ability in dental radiology refers to the skills of dental radiography and patient management in dental radiology. This review found that clinical knowledge and technical skills affected the accuracy of radiographic diagnosis.

Two studies[15,20] considered the influence of clinical knowledge and technical ability on the occurrence of interpretation errors. The study by Milgrom et al[15] found that clinical knowledge and technical ability in dental radiography influenced the occurrence of diagnostic errors. The attributes related to clinical knowledge included poor interpretation and inappropriate use of radiographs. Taking poor quality films and failing to protect the patient were identified as technical abilities affecting the incidence of errors. All these were considered avoidable events.[15] The study by Bussanelli et al[20] discussed the effect of knowledge in a limited manner, focusing only on the extent of knowledge of ICDAS coding and experience using it for caries diagnosis. They found that the understanding of and experience in using the caries classification increased the accuracy of caries diagnosis.

Time pressure

Time pressure is described as the psychological stress resulting from a time shortage to complete tasks.[25] Time pressure has been identified as a contributor to diagnostic errors in medical radiology.[26] Plessas et al [22] investigated the effect of time pressure on the accuracy of the radiographic interpretation of bitewing radiographs. They found that time pressure resulted in lower sensitivity and diagnostic accuracy, leading to an increased occurrence of interpretive errors. None of the other studies considered the effect of time pressure by including a time limit in their study design; therefore, a comparison of the impact of time pressure could not be made in those studies.

Dental training and education programs (location and duration)

Dental training programs vary in content and length in different parts of the world. The location and extent of training were discussed in two papers.[17,18] In the study by Diniz et al,[17] dentists and dental students in Brazil and Switzerland were compared. The two education programs varied in length of training. The authors concluded that the differences in sensitivity of radiographic caries diagnosis between the student groups could be attributed to the differences in training in the two countries. Lazarchik et al[18] studied the effect of the level of training by comparing the accuracy of radiographic caries diagnosis among dentists, first year and senior dental students. They found that increase in years of training (first-year versus senior-year dental students) was associated with improved accuracy of radiographic diagnosis of caries. The remaining studies in this systematic review did not consider the effect of the location and duration of dental training and education program on diagnostic accuracy and the occurrence of interpretation errors. It is, therefore, not possible to conclude if the type and location of dental education influenced the accuracy of radiographic diagnosis.

Clinical experience

Six studies examined the impact of clinical experience on diagnostic accuracy.[16–21] Clinical experience was described as years since graduation[18] or the number of years in clinical practice.[19] The influence of clinical experience on diagnostic accuracy was measured using sensitivity and specificity of radiographic caries diagnosis. Of the six studies, five (over 80 %) found a strong correlation in that clinical experience increased diagnostic specificity but not sensitivity. These studies also found higher sensitivity of radiographic caries diagnosis among participants with less clinical experience (dental students). Overall, these studies found a higher rate of errors in interpretation among less experienced clinicians. However, Bussanelli et al[20] found no correlation between clinical experience and the accuracy of radiographic caries diagnosis. They reported that experienced clinicians were more confident about their diagnosis than novice clinicians. In Mileman et al,[16] the dentist cohort data was historical, and the dentist's clinical experience was not specified.

Cognitive load

Cognitive load has been described as the mental load experienced by a clinician when faced with a challenging situation or a clinical problem. Lack of conceptual knowledge, inefficient reasoning strategies, and task difficulty increase cognitive load.[23,27] Using pupillary response as a measure of cognitive load, Castner et al[19] reported that the cognitive load varied among novice and experienced clinicians depending on panoramic radiographs of different difficulty levels. They found that the cognitive load among students was consistently high regardless of the difficulty of diagnosis (case complexity). However, the cognitive load increased proportionately to the case difficulty among clinicians. The remaining studies in this review did not directly discuss the influence of cognitive load on the occurrence of errors of interpretation.

This systematic review recognised that the factors affecting the interpretation of dental radiographs identified in the included studies are interlinked and have a follow-on effect on each other, as demonstrated in Figure 2. The figure shows that the factors identified in this review either directly or indirectly affect cognitive load, and an overload on the cognitive functions leads to errors of interpretation.

Discussion

In medical radiology, several factors influence the interpretation of radiographs, and failure of these processes leads to errors.[6,8,28] Interpretation errors can

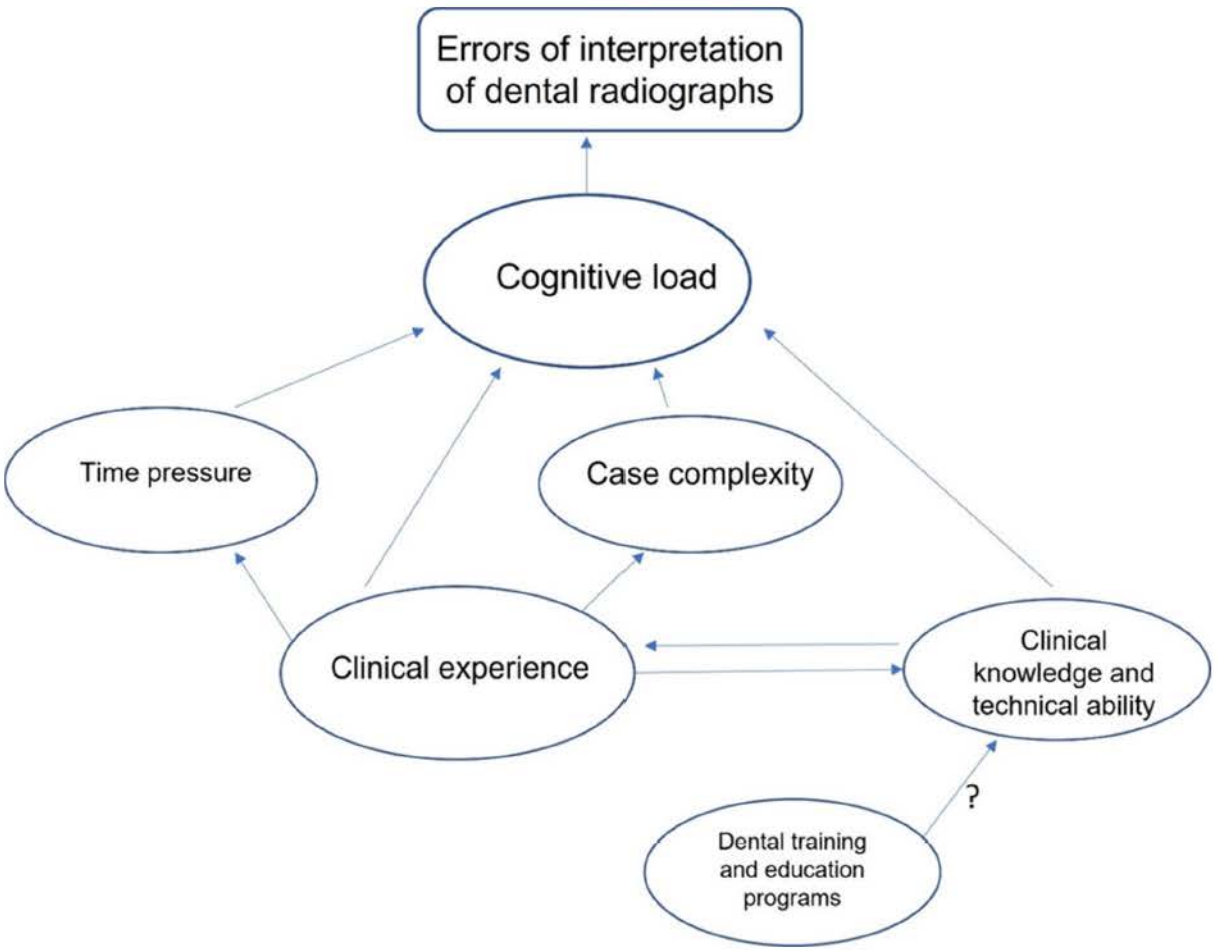


Figure 2 Concept map of factors affecting interpretation of dental radiographs

lead to consequences such as delayed diagnosis of life-threatening diseases, unnecessary or harmful treatment and patient mortality.[4,6,7,11,12,29,30]

The primary goal of this systematic review was to identify from the literature, factors contributing to errors in interpreting dental radiographs. The factors identified from this review included clinical experience, clinical knowledge and technical ability, the complexity of dental radiographs, cognitive load, time pressure and geographic location and length of dental training and education programs. These factors have also been recognised in medical radiology as impacting the accuracy of radiographic diagnosis and interpretive errors.[6,31–34] Research on interpretive errors in medical radiology has shown that interpretive errors are common.[35–37]

This systematic review found that clinical knowledge and technical skills in dental radiography were closely related to clinical experience. Understanding anatomy and pathologic processes, their consequences on oral health and their management are acquired during training, and the knowledge improves over time as a clinician gains more experience.[38] However, this systematic review was inconclusive about how the duration and location of the dental education and training program impacted clinical knowledge and technical skills. Research has shown that experienced clinicians develop clinical *gestalt*, a heuristic decision-making approach.[38–40] The ability to recognise patterns on radiographs and gain a holistic understanding of a radiographic image develops over time. Therefore, experienced clinicians demonstrate diagnostic acuity and higher levels of diagnostic accuracy. Experienced clinicians may also manage time pressure and case complexity differently than novice clinicians. Experienced clinicians are more likely to handle stress (cognitive overload), time pressure and case complexity better.

Several techniques or strategies have been studied in medicine, including templates, checklists, clinical decision support systems and machine learning algorithms to reduce the incidence of errors of interpretation.[41–43] However, in dentistry, there is insufficient data regarding the use of cognitive aids to assist with clinical decision-making.

This systematic review had methodological limitations, including a small number of studies and limiting to papers published only in English. The exclusion of non -English studies may be a source of publication

bias. In addition, publication bias in the included studies could not be assessed due to the heterogeneity of the papers in relation to their study design and significant differences in the effect sizes reported. The quality of evidence was rated as medium because the papers addressed only three of the five criteria described by MMAT. The quality assessment revealed that the study designs did not take appropriate measures to control confounding bias. The studies also had methodological issues, including small sample sizes and inaccurate diagnostic accuracy measurement, which impacted the statistical significance and generalizability of the results. Studies with small sample sizes reduce the power of the study, and the results cannot be extrapolated to a larger population. Diagnostic accuracy studies measure the accuracy of a diagnosis against a gold standard using parameters such as sensitivity and specificity. Studies in this review used different methods to establish gold standard diagnoses, including histopathological sections of teeth and expert panel diagnoses. In one study, gold-standard diagnoses were not used.

Conclusions

This review identified factors affecting the accuracy of radiographic diagnoses in dental radiology. However, it also revealed a notable lack of literature studying the causes and factors influencing errors in interpreting radiographs in a dental setting. All the factors described in this review directly or indirectly impact cognitive load. In turn, cognitive overload can lead to interpretive errors. Errors of interpretation can have a substantial impact on diagnostic and treatment decisions. These errors can potentially affect both the patient and the dental clinician adversely. Well-designed studies in clinical settings are needed to gain insights into how and why interpretive errors occur. Further research will also increase awareness about interpretive errors and pave the way for developing strategies to minimise them in dental practice. Strategies to close these gaps include identifying factors associated with interpretive errors and efforts to build interpretive error epidemiology that will account for the different underlying processes that affect them.

Funding

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

REFERENCES

1. Balogh EP, Miller BT, Ball JR, eds. The National Academies of Sciences, Engineering and M. In: *Improving Diagnosis in Health Care [Internet]*. Washington, DC: The National Academies Press; 2015. Available from: https://www.nap.edu/catalog/21794/improving-diagnosis-in-health-care
2. Corbet EF, Ho DKL, Lai SML. Radiographs in periodontal disease diagnosis and management. *Aust Dent J* 2009; **54 Suppl 1**: S27–43. https://doi.org/10.1111/j.1834-7819.2009.01141.x
3. Fogarty WP, Drummond BK, Brosnan MG. The use of radiography in the diagnosis of oral conditions in children and adolescents. *N Z Dent J* 2015; **111**: 144–50.
4. Krupinski EA. The role of perception in imaging: past and future. *Semin Nucl Med* 2011; **41**: 392–400. https://doi.org/10.1053/j.semnuclmed.2011.05.002
5. Gray JE, Taylor KW, Hobbs BB. Detection accuracy in chest radiography. *AJR Am J Roentgenol* 1978; **131**: 247–53. https://doi.org/10.2214/ajr.131.2.247
6. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive error in radiology. *AJR Am J Roentgenol* 2017; **208**: 739–49. https://doi.org/10.2214/AJR.16.16963
7. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. *Radiographics* 2015; **35**: 1668–76. https://doi.org/10.1148/rg.2015150023
8. Kim, YML. Fool me twice: delayed diagnosis in radiology. *Am J Roentgenol* 2014; **202**: 465–70.
9. Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. *World J Radiol* 2010; **2**: 377–83. https://doi.org/10.4329/wjr.v2.i10.377
10. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. *Radiographics* 2015; **35**: 1668–76. https://doi.org/10.1148/rg.2015150023
11. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 1978; **13**: 175–81. https://doi.org/10.1097/00004424-197805000-00001
12. Berlin L. Accuracy of diagnostic procedures: has it improved over the past five decades? *AJR Am J Roentgenol* 2007; **188**: 1173–78. https://doi.org/10.2214/AJR.06.1270
13. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J* 2021; **372**(7).
14. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. *Annu Rev Public Health* 2014; **35**: 29–45. https://doi.org/10.1146/annurev-publhealth-032013-182440
15. Milgrom P. Quality control of end results: identifying avoidable adverse events in dentistry. *J Am Dent Assoc* 1975; **90**: 1282–90. https://doi.org/10.14219/jada.archive.1975.0279
16. Mileman PA, van den Hout WB. Comparing the accuracy of Dutch dentists and dental students in the radiographic diagnosis of dentinal caries. *Dentomaxillofac Radiol* 2002; **31**: 7–14. https://doi.org/10.1038/sj.dmfr.4600652
17. Diniz MB, Rodrigues JA, Neuhaus KW, Cordeiro RCL, Lussi A. Influence of examiner's clinical experience on the reproducibility and accuracy of radiographic examination in detecting occlusal caries. *Clin Oral Investig* 2010; **14**: 515–23. https://doi.org/10.1007/s00784-009-0323-z
18. Lazarchik DA, Firestone AR, Heaven TJ, Filler SJ, Lussi A. Radiographic evaluation of occlusal caries: effect of training and experience. *Caries Res* 1995; **29**: 355–58. https://doi.org/10.1159/000262092
19. Castner N, Appel T, Eder T, Richter J, Scheiter K, Keutel C, et al. Pupil diameter differentiates expertise in dental radiography visual search. [PLoS One [Internet]]. *PLoS One* 2020; **15**(5): e0223941. https://doi.org/10.1371/journal.pone.0223941

20. Bussaneli DG, Boldieri T, Diniz MB, Rivera LML, Santos-Pinto L, Cordeiro RDCL. Influence of professional experience on detection and treatment decision of occlusal caries lesions in primary teeth. *Int J Paediatr Dent* 2015; **25**: 418–27. https://doi.org/10.1111/ipd.12148
21. Vizzotto MB, Da Silveira PF, Liedke GS, Arus NA, Montagner F, Silveira HLD, et al. Diagnostic reproducibility of the second mesiobuccal canal by CBCT: influence of potential factors. *ORAL Radiol* 2015; **31**: 160–64. https://doi.org/10.1007/s11282-015-0210-z
22. Plessas A, Nasser M, Hanoch Y, O'Brien T, Bernardes Delgado M, Moles D. Impact of time pressure on dentists' diagnostic performance. *J Dent* 2019; **82**: 38–44. https://doi.org/10.1016/j.jdent.2019.01.011
23. Krupinski EA. Current perspectives in medical image perception. *Atten Percept Psychophys* 2010; **72**: 1205–17. https://doi.org/10.3758/APP.72.5.1205
24. Malterud K. The art and science of clinical knowledge: evidence beyond measures and numbers. *Lancet* 2001; **358**: 397–400. https://doi.org/10.1016/S0140-6736(01)05548-9
25. Denovan A, Dagnall N. Development and evaluation of the chronic time pressure inventory. *Front Psychol* 2019; **10**: 2717. https://doi.org/10.3389/fpsyg.2019.02717
26. ALQahtani DA, Rotgans JI, Ahmed NE, Alalwan IA, Magzoub MEM. The influence of time pressure and case complexity on physicians' diagnostic performance. *Health Professions Education* 2016; **2**: 99–105. https://doi.org/10.1016/j.hpe.2016.01.006
27. Iskander M. Burnout, cognitive overload, and metacognition in medicine. *Med Sci Educ* 2019; **29**: 325–28. https://doi.org/10.1007/s40670-018-00654-5
28. Pinto A, Brunese L, Pinto F, Reali R, Daniele S, Romano L. The concept of error and malpractice in radiology. *Semin Ultrasound CT MR* 2012; **33**: 275–79. https://doi.org/10.1053/j.sult.2012.01.009
29. Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. *Ulster Med J* 2012; **81**: 3–9.
30. Waite S, Grigorian A, Alexander RG, Macknik SL, Carrasco M, Heeger DJ, et al. Analysis of perceptual expertise in radiology - current knowledge and a new perspective. *Front Hum Neurosci* 2019; **13**: 213. https://doi.org/10.3389/fnhum.2019.00213
31. Renfrew DL, Franken EAJ, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. *Radiology* 1992; **183**: 145–50. https://doi.org/10.1148/radiology.183.1.1549661
32. Pinto A, Acampora C, Pinto F, Kourdioukova E, Romano L, Verstraete K. Learning from diagnostic errors: a good way to improve education in radiology. *Eur J Radiol* 2011; **78**: 372–76. https://doi.org/10.1016/j.ejrad.2010.12.028
33. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. *AJR Am J Roentgenol* 2013; **201**: 611–17. https://doi.org/10.2214/AJR.12.10375
34. Schwendicke F, Tzschoppe M, Paris S. Radiographic caries detection: a systematic review and meta-analysis. *J Dent* 2015; **43**: 924–33. https://doi.org/10.1016/j.jdent.2015.02.009
35. Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging* 2017; **8**: 171–82. https://doi.org/10.1007/s13244-016-0534-1
36. Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual and interpretive error in diagnostic radiology-causes and potential solutions. *Acad Radiol* 2019; **26**: 833–45. https://doi.org/10.1016/j.acra.2018.11.006
37. Berlin L. Radiologic errors, past, present and future. *Diagnosis (Berl)* 2014; **1**: /j/dx.2014.1.issue-1/dx-2013-0012/dx-2013-0012.xml: 79–84: . https://doi.org/10.1515/dx-2013-0012
38. Preisz A. Fast and slow thinking; and the problem of conflating clinical Reasoning and ethical deliberation in acute decision-making. *J Paediatr Child Health* 2019; **55**: 621–24. https://doi.org/10.1111/jpc.14447
39. Cervellin G, Borghi L, Lippi G. Do clinicians decide relying primarily on bayesians principles or on gestalt perception? some pearls and pitfalls of gestalt perception in medicine. *Intern Emerg Med* 2014; **9**: 513–19. https://doi.org/10.1007/s11739-014-1049-8
40. Koontz NA, Gunderman RB. Gestalt theory: implications for radiology education. *AJR Am J Roentgenol* 2008; **190**: 1156–60. https://doi.org/10.2214/AJR.07.3268
41. Berenholtz SM, Pronovost PJ, Lipsett PA, Hobson D, Earsing K, Farley JE, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. *Crit Care Med* 2004; **32**: 2014–20. https://doi.org/10.1097/01.ccm.0000142399.70913.2f
42. Weiss MJ, Kramer C, Tremblay S, Côté L. Attitudes of pediatric intensive care unit physicians towards the use of cognitive aids: a qualitative study. *BMC Med Inform Decis Mak* 2016; **16**: 53. https://doi.org/10.1186/s12911-016-0291-6
43. Krombach JW, Edwards WA, Marks JD, Radke OC. Checklists and other cognitive AIDS for emergency and routine anesthesia care-A survey on the perception of anesthesia providers from a large academic us institution. *Anesth Pain Med* 2015; **5**(4): e26300. https://doi.org/10.5812/aamp.26300v2

Chapter 3

Australian Dentist's Knowledge and Perceptions of Factors Affecting Radiographic Interpretation

Gathering clinicians' perceptions through surveys offers an overview of the clinicians' understanding and perceptions about interpretive errors, highlighting both known and unknown aspects in this important area. This chapter presents a survey of dental practitioners conducted to investigate the perceptions about factors contributing to interpretive errors.

Scientific Research Report

Australian Dentist's Knowledge and Perceptions of Factors Affecting Radiographic Interpretation

Shwetha Hegde [a*], *Jinlong Gao* [b], *Rajesh Vasa* [c], *Shanika Nanayakkara* [b], *Stephen Cox* [a]

[a] *Sydney Dental School, University of Sydney, Surry Hills, NSW, Australia*
[b] *Institute of Dental Research, Westmead Centre for Oral Health, University of Sydney, Westmead, NSW, Australia*
[c] *Applied Artificial Intelligence, Deakin University, Melbourne, Australia*

ARTICLE INFO

Article history:
Received 17 August 2023
Received in revised form
3 November 2023
Accepted 5 November 2023
Available online 6 January 2024

Key words:
Cognition
Dental radiology
Interpretive errors
Perception

ABSTRACT

Background: Errors of interpretation of radigraphic images, also known as interpretive errors, are a critical concern as they can have profound implications for clinical decision making. Different types of interpretive errors, including errors of omission and misdiagnosis, have been described in the literature. These errors can lead to unnecessary or harmful treat/or prolonged patient care. Understanding the nature and contributing factors of interpretive errors is important in developing solutions to minimise interpretive errors. By exploring the knowledge and perceptions of dental practitioners, this study aimed to shed light on the current understanding of interpretive errors in dentistry.

Methods: An anonymised online questionnaire was sent to dental practitioners in New South Wales (NSW) between September 2020 and March 2022. A total of 80 valid responses were received and analysed. Descriptive statistics and bivariate analysis were used to analyse the data.

Results: The study found that participants commonly reported interpretive errors as occurring 'occasionally', with errors of omission being the most frequently encountered type. Participants identi ed several factors that most likely contribute to interpretive errors, including reading a poor-quality image, lack of clinical experience and knowledge, and excessive workload. Additionally, general practitioners and specialists held different views regarding factors affecting interpretive errors.

Conclusion: The survey results indicate that dental practitioners are aware of the common factors associated with interpretive errors. Errors of omission were identi ed as the most common type of error to occur in clinical practice. The ndings suggest that interpretive errors result from a mental overload caused by factors associated with image quality, clinician-related, and image interpretation. Managing and identifying solutions to mitigate these factors are crucial for ensuring accurate and timely radiographic diagnoses. The ndings of this study can serve as a foundation for future research and the development of targeted interventions to enhance the accuracy of radiographic interpretations in dentistry.

© 2023 The Authors. Published by Elsevier Inc. on behalf of FDI World Dental Federation.
This is an open access article under the CC BY license
(http://creativecommons.org/licenses/by/4.0/)

Introduction

The interpretation of radiographs involves 2 identi able and nonseparable components of diagnosis: identifying the abnormality on a radiograph (perception) and recognising the signi cance of the abnormal nding (cognition).[1,2] These 2 components usually function effectively; however, they are not awless, and errors can result from shortcomings in perception and cognition. Errors of interpretation, or interpretive

* Corresponding author. Sydney Dental School, University of Sydney, Rm 7.20, Level 7, 2 Chalmers St, Surry Hills, NSW 2010 Australia
E-mail address: shwetha.hegde@sydney.edu.au (S. Hegde).
Shwetha Hegde: http://orcid.org/0000-0003-1429-9149
Rajesh Vasa: http://orcid.org/0000-0003-4805-1467
Shanika Nanayakkara: http://orcid.org/0000-0002-5479-9650
https://doi.org/10.1016/j.identj.2023.11.006
0020-6539/© 2023 The Authors. Published by Elsevier Inc. on behalf of FDI World Dental Federation. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

errors, are a type of diagnostic error and have been described as incorrect interpretation[3] or a discrepancy in the interpretation that signi cantly differs from the consensus among clinicians.[1] Diagnostic errors in radiology can lead to misdiagnosis and inappropriate treatment, compromising safe and high-quality patient care. The consequences of diagnostic errors vary in severity and depend on the type of error, the stage of the diagnostic process, and the time taken to discover and rectify them. Errors may cause direct or indirect, permanent or temporary harm to the patient or may not result in any harm.[1,4,5] Since Leo Garland's pioneering work in 1949,[6] several studies have described the error types and their frequency in medical radiology.[3,7,8] Medical radiology error rates have remained more or less unchanged since Garland's report.[9] Currently, the global frequency of interpretive error ranges from 3% to 20%. The prevalence of interpretive errors can vary depending on the different regions worldwide, level of expertise and speci c eld within health care.[1,3,10,11]

Interpretive errors have been classi ed into 4 types based on the underlying processes: [1,2,4,8,12-14] near misses where the error is detected and corrected before the patient is harmed,[15] delayed diagnoses where a diagnosis is not made within a reasonable amount of time,[16] misdiagnoses where incorrect or wrong diagnosis occurs[16] and errors of omission when the diagnosis or nding is missed on a radiograph.[16] Factors contributing to interpretive errors in radiology are circumstances and actions that lead to or increase the risk of an interpretative error.[17] Several factors contributing to interpretive errors in radiology have been identi ed in the literature, including clinical experience, clinical knowledge, clinician fatigue and burnout.[3,8,12,18-21] These factors are described under 4 domains, including image interpretation-related, image quality-related, clinician performance-related and patient-related factors in line with the WHO's International Classi cation of Patient Safety report.[17] This classi cation has been used in the context of medical imaging,[22,23] and surgical care in organ transplants.[24] Image interpretation-related factors involve cognitive and perceptual processes, which can result in false positive and false negative diagnoses. Image-related factors pertain to aspects of image quality, including lm quality and lm faults. Clinician-related factors encompass factors such as lack of knowledge and clinical experience and ill health. Patient-related factors refer to situations such as patients experiencing pain, patients with dental fear and demanding patients.

Interpretive errors have been analysed using self-reported data, including surveys, clinical audits and autopsy results.[25] Because of the dynamic nature of the diagnostic process, it is challenging to collect real-time data on interpretive errors except in experimental settings. In addition, data gathered in experimental settings cannot always be translated into real-life situations. In some cases where a de nitive diagnosis is not possible based on a radiograph, 'errors' are just a difference of opinion rather than an incorrect diagnosis, and a clear-cut distinction between error and observer variation cannot be made.[3,26] For these reasons, studying the frequency and underlying causes of errors can be challenging. Obtaining clinicians' perspectives through self-reported surveys that assess the knowledge and attitudes provides a snapshot of their understanding and perception of interpretative errors and highlights what is known and unknown in this area.[27]

While there is extensive data on this topic in the medical literature,[13,26,28-30] there is limited understanding of how and why interpretive errors occur in dentistry. In addition, it is unclear whether dental professionals make the same mistakes or miss similar ndings while analysing dental radiographs and whether the same factors in uence their decision-making. Dentists routinely take and interpret radiographs, whereas, in a medical setting, radiographs are sent to a radiologist for interpretation and diagnosis. This distinction may also in uence the incidence of errors of interpretation and its impact on patient care. A few studies have assessed the in uence of clinical experience, stress and time pressure on the accuracy of radiographic diagnosis of certain common dental conditions.[31-34]

Our recent systematic review identi ed factors affecting the interpretation of dental radiographs, including clinical experience, time pressure, clinical knowledge, case complexity and cognitive load.[35] Understanding the causes and factors affecting interpretive errors will inform the development of solutions to minimise errors in clinical practice, improve diagnostic performance, and enhance patient outcomes. This survey aimed to evaluate dental practitioners' knowledge and perceptions about errors of interpretation of dental radiographs and the factors affecting them.

Method

The study was approved by the Human Research Ethics Committee at the University of Sydney (2020/336). The survey was an online anonymised questionnaire, and informed consent was assumed when the participants completed and submitted the survey. Registered dentists and oral health therapists in NSW, Australia, were invited to participate in the survey. The survey was distributed via email and was also posted on social media and the ADA's e-newsletter.

The survey was carried out in 2 rounds between September 2020 and July 2021 and again between January 2022 and March 2022. The second round was conducted to boost the response rate. During this period, the data collected was monitored regularly for response patterns. A manual assessment was made to determine that data saturation was reached, as no new information was gathered with the new responses received similar to the data saturation approach[36] used in qualitative research. Considering the voluntary nature of the survey, multiple survey platforms such as emails, social media and ADA's newsletter were used to improve coverage and response rate.

The online questionnaire was created using the Qualtrics platform (www.qualtrics.com), which was also used to distribute the survey to collect data. The questionnaire consisted of both closed and open-ended questions and was self-administered. Since no existing questionnaire in the dental and medical literature could be adapted to this study, the questions were based on the review of relevant literature on clinical decision-making, interpretive errors,[35] and the clinical experience of the research team. Ten dentists were invited to participate in the pilot study to ensure the relevance and appropriateness of the survey questions. Two experts in survey methods and radiological

diagnosis were also consulted to determine the content validity of the questionnaire. The pilot study and expert consultation helped re ne and improve the quality of the survey questions and ensured that the nal questionnaire was valid and reliable.

The questionnaire (Supplementary File Questionnaire) was organised into the following sections. The rst part assessed the participants' awareness of errors in interpreting radiographs in their clinical practice. The second section explored the perception of the frequency of errors of interpretation. The third section explored the participants' opinions on the factors contributing to errors. The last part of the questionnaire collected participants' demographic characteristics.

Data from the survey portal www.qualtrics.com was exported and analysed using IBM SPSS statistics software (Version 26, IBM SPSS Inc). Cronbach's alpha measured the internal consistency of the survey. Descriptive statistics were used to summarise and present the demographic characteristics of the participants and responses. Questions, where the respondents were required to rank the choices, were summarised using average weighted scores. Weights were applied in reverse, providing the highest score for the rst choice and the lowest score for the last choice. The association between the categorical variables was assessed using the Chi-square test of independence (or Fisher's exact test when more than 20% of cells had expected frequencies <5). The Spearman rank correlation test assessed the correlation among the variables. The signi cance level for all statistical tests was set at P-value < .05.

Results

One hundred eighteen dental practitioners responded to the survey. After removing the records with incomplete data, responses from 80 participants were included in the analysis (67.8%). Cronbach's alpha value for the questionnaire was 0.928, indicating a high internal consistency for the survey instrument.

Demographic characteristics of participants

As presented in Table 1, the age of the participating dental practitioners ranged from 23 to 72 years, with 57.5% self-identi ed as females. The participants' clinical experience ranged from 1 to 50 years, with most participants in the 10 to 19 years' experience range. As shown in Table 1, on average, the participants reported requesting radiographs for 6 patients per day, taking an average of 4 periapical radiographs, 5 bitewing radiographs and 3 orthopantomograms (OPGs) per day. The survey participants were employed in private practice, public practice, academia, or a combination of all three. (Table 1).

Knowledge of the incidence and management of interpretive errors in clinics

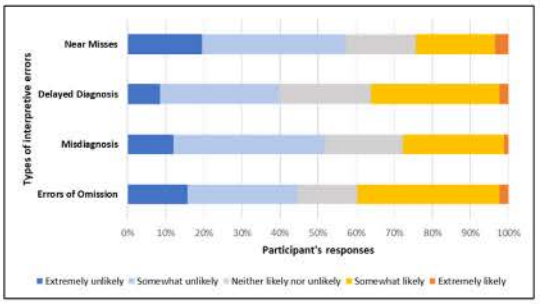
Responses on how often interpretive errors occurred in their clinical practice were classi ed into 4 categories: regularly, occasionally, rarely and unknown, with the most common response category being 'occasionally'. The majority of the participants (n = 49, 61.3%) indicated that interpretive errors were documented in their clinical practice. In addition, the majority of them (n = 61, 76.3%) indicated that their patients were informed about the error when it occurred. The survey found that a majority of the participants (n = 48, 60.8%) selected 'sometimes' to describe the extent to which errors were preventable. Interestingly, statistically signi cant (P-value = .02) differences were found in the responses of general practitioners (GP) and specialist dentists. More GPs (n = 12, 35%) thought that errors were preventable (always/most of the time) than specialists (n = 2, 14%).

Table 1 – Demographic characteristics of the study participants.

Characteristics	Value (n = 80)
Age (years)	
Mean (SD)	41.7±11.0
Minimum	23
Maximum	72
Gender n (%)	
Female	46 (57.5)
Male	22 (27.5)
Not speci ed	12 (15.0)
Type of clinical practice (choice count) n (%)	
Type of practice*	
General	35 (43.7)
Specialist	14 (17.5)
Not speci ed	31(38.8)
Sector†	
Private	21 (26.3)
Public	14 (17.5)
Both	11 (13.8)
Practice location†	
Metropolitan	13 (16.4)
Rural	5 (6.3)
Other (university academic)	2 (0.03)
Clinical experience in years	
Mean (SD)	17.1 (±11.8)
Minimum	1
Maximum	50
Time spent doing different activities per week (mean hours (SD))	
Clinical	27.90(±12.5)
Research	3.09 (±5.4)
Teaching	5.76 (±7.5)
Administration	4.83 (±5.9)
Patients per day requiring radiographs (mean ± SD)	6.2 (± 2.5)
Number of radiographs taken per day (mean± SD)	
Periapical	3.8 (±3.4)
Bitewings	4.9 (±3.7)
Panoramic radiographs (OPG)	2.7 (±2.7)
Other	0.49 (±1.04)
Imaging systems used in clinical practice n (%) (Choice count)†	
Photostimulable phosphor (PSP)	42 (52.5)
Direct digital (CCD, CMOS)	25 (31)
Analog lms (chemically processed)	14 (17.5)
Other	20 (25)

* The total number of responses does not add up to 100% for some variables due to missing responses.
† As the participants have selected more than one choice, the responses don't add up to 100%.

A



A: Types of interpretive errors and their perceived frequency
B: Perceptions about the occurrence of four types of interpretive errors among GP and specialist dentists

B

Type of error	Participants qualifications n (%)	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Extremely likely	p-value*
Omission error	General Dentist (n=35)	3(8.5)	8(22.8)	9(25.7)	15(42.8)	0	**0.045**
	Specialist (n=14)	2(14)	8(57)	0	4(28.5)	0	
Delayed diagnosis	General Dentist (n=35)	2(5.7)	9(25.7)	14(40)	10(28.5)	0	**0.032**
	Specialist (n=14)	1(7)	7(50)	1(7)	3(21)	2(14)	
Misdiagnosis	General dentist (n=35)	6 (17)	12 (33.6)	9 (25.7)	7(19.6)	1 (2.8)	0.790
	Specialist (n=14)	1(7)	6(42)	3(21)	4(28)	0	
Near misses	General dentist (n=35)	5(14)	13(36.4)	10(28)	6(16.8)	1 (2.8)	0.736
	Specialist (n=14	3(21)	7(50)	2(14)	2(14)	0	

Note: The total responses do not add up to n=80 as responses were not available for some participants.
*Statistically significant values are in bold

Fig. 1 – Composite figure showing the perceived frequency of interpretive errors and differences in the opinions of general practitioners (GPs) and specialists.

Types of interpretive errors and dental conditions frequently associated with interpretive errors

The participants acknowledged that all 4 types of errors could occur in clinical practice. As shown in Figure 1, errors of omission (n = 31, 39.2%) were identi ed as the most frequent type of error, followed by delayed diagnosis (n = 29, 36.7%), misdiagnosis (n = 22, 27.9%), and near-miss errors (n = 21, 26.6%). There were no statistically signi cant differences in the responses among different demographic categories, such as gender and location of practice, except in the responses given by general GPs and specialists. GPs suggested that omission errors occurred more frequently, while specialists suggested that delayed diagnosis occurred more frequently. In addition, the participants ranked dental caries, cracked teeth and cervical resorption of teeth as the most common dental conditions associated with interpretive errors (Figure 2).

Factors contributing to interpretive errors

Twenty-two factors affecting interpretation errors are categorised into 4 domains, and the results are presented in Figure 3. The ratings are on a 5-point Likert scale ranging from extremely likely to extremely unlikely. The top 5 factors ranked most likely: poor image quality, lack of clinical experience, excessive workload, jumping to conclusions, and misdiagnosis due to a

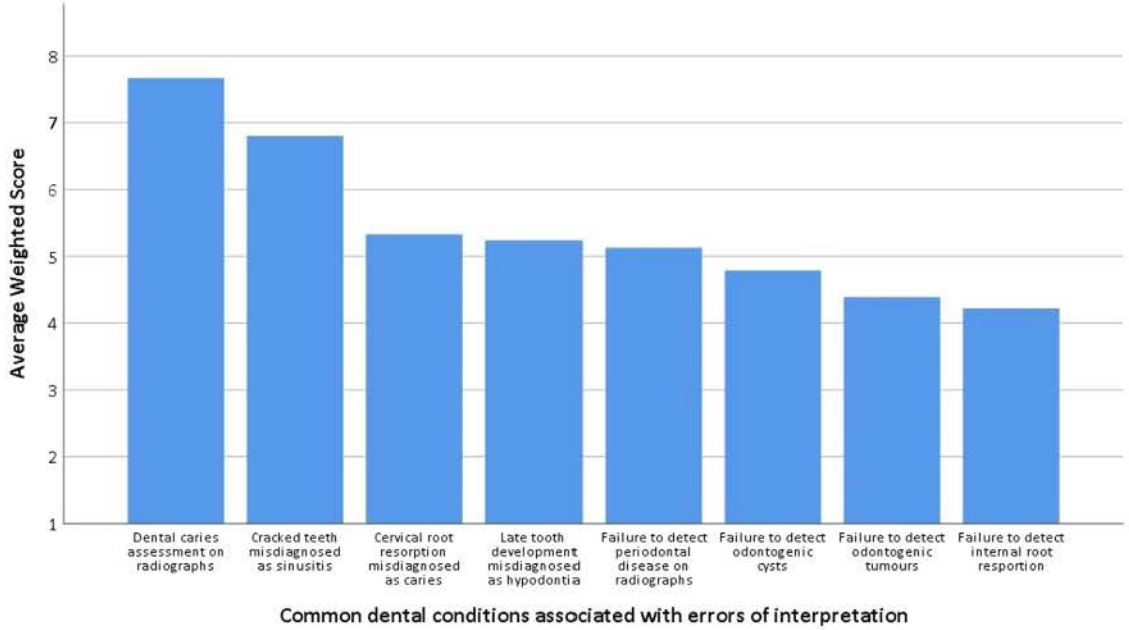


Fig. 2 – Rank scores of participants' perceptions of the likelihood of common oral conditions being incorrectly diagnosed on radiographs.

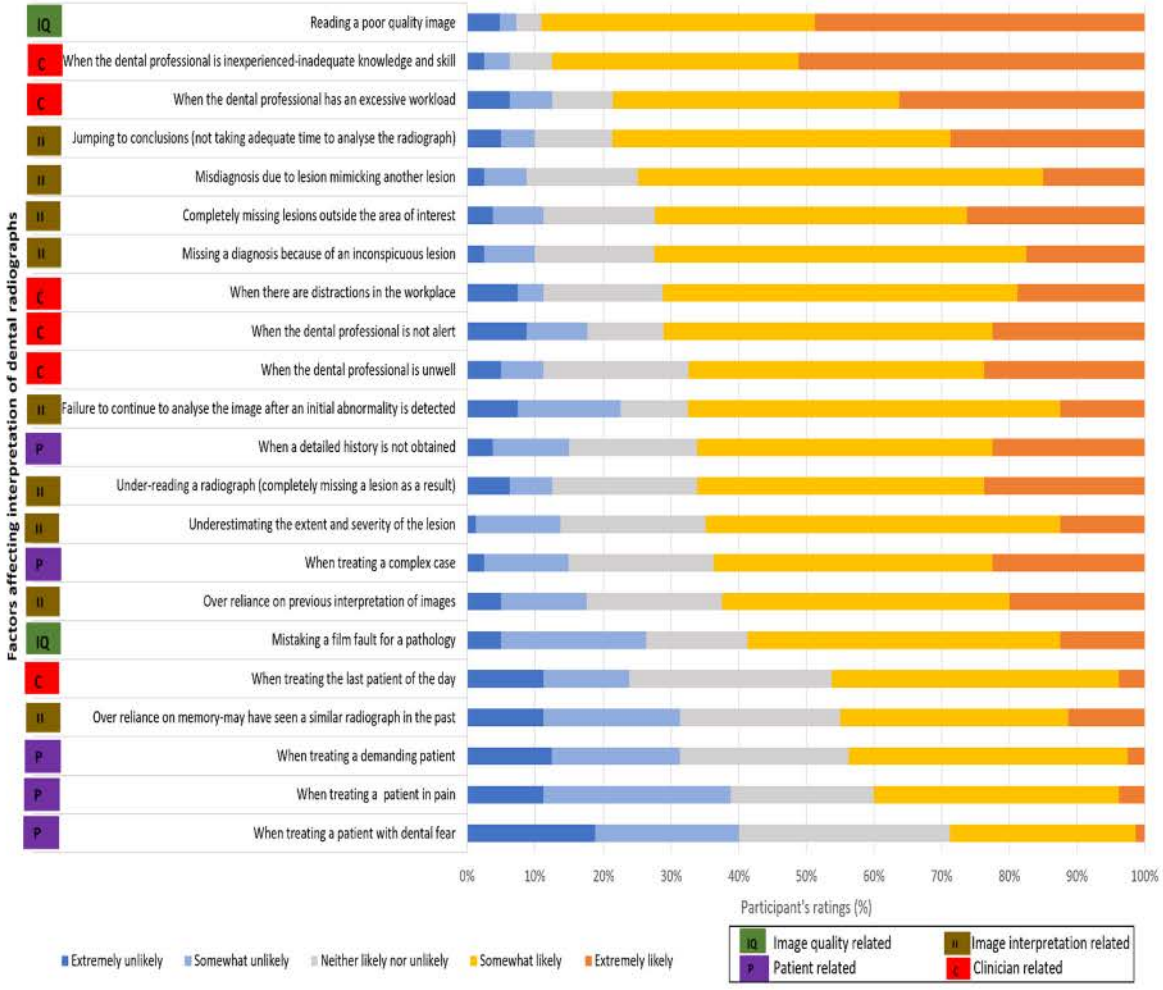


Fig. 3 – Participants' responses to the likelihood of factors affecting image interpretation. The factors have been listed in descending order from most likely to least likely as identified by the survey participants.

The figure also uses a colour coding system to identify the factors affecting interpretive errors based on the 4 domains.

lesion mimicking another lesion, while the 5 least likely factors identi ed were treating the last patient of the day over-reliance on memory, treating a demanding patient, treating a patient in pain, and treating a patient with dental fear.

More GPs thought a lack of alertness and poor patient history could lead to interpretive errors (Table 2). In contrast, more specialists indicated that treating the last patient of the day and overly relying on memory could lead to interpretive

Table 2 – Differences in the responses between different participant groups regarding the likelihood of occurrence of factors affecting interpretation. Only statistically significant responses are shown.

	Participant's quali cations N (%)	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Extremely likely	P-value*
Dentist not alert	General Dentist (n = 35)	3 (8.5)	2 (5.7)	4 (11)	20 (57)	6 (17)	**0.025**
	Specialist (n = 14)	2 (14)	2 (14)	2 (14)	1 (7)	7 (50)	
Treating the last patient of the day	General Dentist (n = 35)	3 (8.5)	4 (11)	14 (40)	14 (40)	0	**0.039**
	Specialist (n = 14)	4 (28.5)	0	4 (28.5)	4 (28.5)	2 (14)	
Poor patient history	General Dentist (n = 35)	2 (5.7)	2 (5.7)	8 (22.8)	16 (45.7)	7 (20)	**0.046**
	Specialist (n = 14)	1 (7)	5 (35.7)	2 (14)	2 (14)	4 (28)	
Over-reliance on memory	General Dentist (n = 35)	4 (11)	8 (22.8)	12 (34)	10 (28.5)	1 (2.8)	**0.034**
	Specialist (n = 14)	1 (7)	2 (14)	1 (7)	6 (42.8)	4 (28.5)	

* Statistically signi cant results.

errors. However, it is important to note that this result may be subject to bias since the sample size of specialists was small (n = 14). In addition, the survey found that more female respondents (n = 24, 52%) said that treating the last patient was extremely likely/somewhat likely to lead to interpretive errors than male respondents (n = 7, 31.8%) (Statistically signi cant; P-value = .046). However, there were no statistically signi cant differences in the responses of other demographic categories (GP vs specialists or dental practitioners in private and public dental practices).

The correlation coef cients between various factors that affect interpretation errors in radiographic images are presented in a table in the Supplementary Table S1. A signi cant positive correlation was found between over-reliance on previous interpretations and jumping to conclusions (P-value <.001). Several factors were found to have moderate positive correlations with the occurrence of interpretive errors, such as excessive workload, being the last patient of the day, workplace distraction, inexperienced dentists, patients with dental fear and absence of dental history (P-value <.001). Mistaking lm fault for pathology had a moderate positive correlation (P-value <.001) with errors of interpretation, suggesting the signi cance of high-quality radiographs in diagnosis and the need for proper training and attention to technical aspects of radiography. On the other hand, a signi - cant negative correlation was observed between clinical experience and perception of case complexity as a cause of the interpretive error (-0.258) (P-value <.05), indicating that more experienced clinicians were less likely to make errors when treating patients with complex dental problems.

Discussion

The purpose of the survey was to evaluate the knowledge and perceptions of dental practitioners regarding errors in dental radiographic interpretations. The survey ndings revealed that dental professionals considered errors of omission as the most likely type of error to occur compared to misdiagnosis, near misses and delayed diagnosis. In contrast, the medical literature suggests that delayed diagnosis and misdiagnosis are more likely to occur.[8,30,37] In medicine and dentistry, errors of omission occur when a clinician fails to identify a pathology on a radiograph. They are more likely to occur in dental practice due to the setting where radiographic interpretation occurs, ie, in a busy dental clinic, compared to a dedicated image reading room in medical radiology. Hence, some radiographic ndings could be missed. Additionally, there was a difference in the perceptions of GPs and dental specialists, where GPs believed that omission errors were more likely to occur, while specialists believed that delayed diagnosis occurred more frequently. This may be because specialists see more complex cases referred by GPs.

Dental caries, cracked teeth, and cervical resorption of teeth were identi ed as the most common dental conditions associated with interpretive errors, suggesting that these conditions pose particular challenges in radiographic diagnosis. Our ndings are in agreement with several studies that have demonstrated that radiographic assessment of caries depth and cracked teeth were more likely to be misdiagnosed.[38-42] The skills to detect and diagnose caries and cracked teeth develop as clinicians gain more experience. Moreover, the survey responses also showed that interpretation errors were less likely to occur with diseases such as odontogenic cysts and tumours. These conditions are encountered less frequently, so dental practitioners are likely to be more cautious and attentive during the diagnostic process.

The practice of dentistry requires a high level of attention and places a high demand on perception and cognitive functions. Factors such as the clinician's physical health and alertness, time of day, and workplace distractions, have been shown to in uence cognitive and perceptual functions.[43,44] Studies have shown that physical and mental fatigue can lead to interpretive errors among radiologists and are a well-known cause of such errors in medical radiology.[45] In this survey, when the factors contributing to interpretive errors were ranked in the order of likelihood, the image quality-related, clinician-related, and image interpretation-related were prominent within the 50th percentile. In contrast, the patient-related factors dominated the 50th to 100th percentiles. Poor radiographic image quality can signi cantly hinder radiographic diagnosis by making it dif cult or impossible for clinicians to visualise and accurately interpret the relevant anatomical structures or lesions. In addition, the participants also recognised the signi cance of clinician-related and image-interpretation-related factors in the occurrence of errors of interpretation. Our previous systematic review[35] has identi ed that these factors lead to mental overload resulting in interpretive errors. Mental overload can be managed with cognitive aids such as checklists and machine learning supporting clinical decision support, as demonstrated in the medical radiology literature.[46-49]

It is interesting to note that the participants did not perceive patient-related factors as signi cantly impacting the frequency of interpretive errors. This observation suggests that dental practitioners do not view these factors as signi - cant contributors to interpretive errors. The in uence of a patient's physical and mental state on a clinician's decision-making has been studied, revealing that patient behaviour, motivation for self-care, the patient-clinician relationship and the patient's mental health can affect the clinical decisions made by clinicians.[44,50,51]

Limitations

This study had a limited sample size, with the majority of the participants practising in urban locations, limiting the generalisability of the results. Similar small sample sizes were observed in other survey-based studies conducted in Australian dental communities in the recent past.[52-54] Furthermore, the voluntary nature of the survey could also result in selection bias. Although the open invitation to the survey was sent out to all dental practitioners, responses were not received from other dental practitioners, such as oral health therapists, who are also involved in the dental radiography and interpretation process. However, the survey population was representative when compared to the wider dental

community using the most recent registrant data (Dental Board of Australia publication June 2023[55]). In addition, the inclusion of a proportionate number of specialists, especially oral radiologists, would have been bene cial in obtaining a more comprehensive understanding of factors affecting interpretive errors.

Conclusion

The survey on the knowledge and perceptions regarding interpretive errors in dentistry provided valuable insights into the prevalence and factors affecting interpretive errors. The ndings emphasised the need for careful attention to detail and continuous education in radiographic interpretation, especially in identifying errors of omission, which were found to be the most common. The study also highlighted the importance of various factors related to image quality, clinicians, and image interpretation which may contribute to interpretive errors. To reduce interpretive errors, the use of radiographic diagnostic aids, such as computerised decision support systems and machine learning algorithms and identifying potential solutions to minimise cognitive load should be considered. Future studies in collaboration with dental organisations should focus on gaining a deeper understanding of the pattern of image analysis by dental practitioners and identifying new insights into factors in uencing their occurrence. Qualitative studies including focus group interviews, may also help in an in-depth understanding of the attitudes and perceptions of dental practitioners regarding interpretive errors.

Conflict of interest

The authors declare that they have no known competing nancial interests or personal relationships that could have appeared to in uence the work reported in this paper.

Author contributions

Substantial contribution to the conception, design, acquisition, and analysis of data; drafting and revising the manuscript: Shwetha Hegde.

Review of study design, feedback on the concept, data analysis and manuscript preparation: Jinlong Gao.

Review of study design, feedback on the concept, data analysis and manuscript preparation: Stephen Cox.

Data analysis and manuscript preparation: Shanika Nanayakkara.

Review of study design, feedback on concept, manuscript review & refinement: Rajesh Vasa.

Funding

This project was partly supported by the Australian Dental Research Fund (grant number 2790/2020).

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.identj.2023.11.006.

REFERENCES

1. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. Radiographics [Internet] 2015;35(6):1668–76. doi: 10.1148/rg.2015150023.
2. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol 1978;13(3):175–81.
3. Waite S, Scott J, Gale B, Fuchs T, Kolla S, Reede D. Interpretive error in radiology. AJR Am J Roentgenol 2017;208(4):739–49.
4. Onder O, Yarasir Y, Azizova A, Durhan G, Onur MR, Ariyurek OM. Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review. Insights Imaging [Internet] 2021;12(1):51. doi: 10.1186/s13244-021-00986-8.
5. Caranci F, Romano L, Carra ello G, Fonio P, Brunese L Learning from errors in radiology: a comprehensive review. Semin Ultrasound, CT MRI [Internet] 2012;33(4):379–382. [cited 2018 December 5]. Available from: https://www.sciencedirect.com/science/article/pii/S0887217112000169. Accessed 14 June 2022.
6. Garland HL. On the scienti c evaluation of diagnostic procedures. Radiology 1949;52(3):309–28.
7. Scott IA, Crock C. Diagnostic error: incidence, impacts, causes and preventive strategies. Med J Aust 2020;213(7):302–305.e2.
8. Kim YW ML. Fool me twice: delayed diagnosis in radiology. Am J Roentgenol 2014;202:465–70.
9. Berlin L Radiologic errors, past, present and future. Diagnosis [Internet] 2014;1(1):79–84. Available from: https://www.degruyter.com/view/j/dx.2014.1.issue-1/dx-2013-0012/dx-2013-0012.xml. Accessed 30 June 2022.
10. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. BMJ Qual Saf 2017;26(6):484–94.
11. Graber ML, Kissam S, Payne VL, et al. Cognitive interventions to reduce diagnostic error: a narrative review. BMJ Qual Saf 2012;21(7):535–57.
12. Krupinski EA. The role of perception in imaging: past and future. Semin Nucl Med [Internet] 2011;41(6):392–400. doi: 10.1053/j.semnuclmed.2011.05.002.
13. Waite S, Scott JM, Legasto A, Kolla S, Gale B, Krupinski EA. Systemic error in radiology. Am J Roentgenol 2017;209(3):629–39.
14. Pinto A, Acampora C, Pinto F, Kourdioukova E, Romano L, Verstraete K. Learning from diagnostic errors: a good way to improve education in radiology. Eur J Radiol [Internet] 2011;78(3):372–6.
15. Thornton RH, Miransky J, Killen AR, Solomon SB, Brody LA. Analysis and prioritization of near-miss adverse events in a radiology department. Am J Roentgenol 2011;196(5):1120–4.
16. Newman-Toker DE. A uni ed conceptual model for diagnostic errors: underdiagnosis, overdiagnosis, and misdiagnosis. Diagnosis (Berlin, Ger.) 2014;1(1):43–8.
17. WHO. Conceptual Framework for International Classi cation for Patient Safety. 2009.
18. Iskander M. Burnout, cognitive overload, and metacognition in medicine. Med Sci Educ [Internet] 2019;29(1):325–8. doi: 10.1007/s40670-018-00654-5.
19. Preisz A. Fast and slow thinking; and the problem of con ating clinical reasoning and ethical deliberation in acute decision-making. J Paediatr Child Health [Internet] 2019;55(6):621–4.

20. Patel SH, Itri JN. The role of intuitive cognition in radiologic decision making. J Am Coll Radiol [Internet] 2022;19(5):669–76.
21. Weber LS, Schulze R. The effects of examiner fatigue on the diagnostic accuracy of dental radiographs. Clin Oral Investig 2021;25(11):6193–9.
22. Jones DN, Thomas MJW, Mandel CJ, et al. Where failures occur in the imaging care cycle: lessons from the radiology events register. J Am Coll Radiol [Internet] 2010;7(8):593–602.
23. Kasalak Ö, Yakar D, Dierckx RA, Kwee TC. Patient safety incidents in radiology: frequency and distribution of incident types. Acta Radiol 2021;62(5):653–66.
24. McElroy LM, Woods DM, Yanes AF, et al. Applying the WHO conceptual framework for the International Classi cation for Patient Safety to a surgical population. Int J Qual Heal Care 2016;28(2):166–74.
25. Balogh EP, Miller BT, Ball JR, The National Academies of Sciences, Engineering. Improving diagnosis in health care. Washington, DC: The National Academies Press; 2015. p. 1–472.
26. Robinson PJ. Radiology's Achilles' heel: error and variation in the interpretation of the Röntgen image. Br J Radiol [Internet] 1997;70(839):1085–98. doi: 10.1259/bjr.70.839.9536897.
27. Kelley K, Clark B, Brown V, Sitzia J. Good practice in the conduct and reporting of survey research. Int J Qual Heal Care [Internet] 2003;15(3):261–6. doi: 10.1093/intqhc/mzg031.
28. Sabih De, Sabih A, Sabih Q, Khan AN. Image perception and interpretation of abnormalities; can we believe our eyes? Can we do something about it? Insights Imaging [Internet] 2011;2(1):47–55. doi: 10.1007/s13244-010-0048-1.
29. Busby LP, Courtier JL, Glastonbury CM. Bias in radiology: the how and why of misses and misinterpretations. Radiogr a Rev Publ Radiol Soc North Am Inc 2018;38(1):236–47.
30. Renfrew DL, Franken EAJ, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classi cation and lessons in 182 cases presented at a problem case conference. Radiology 1992;183(1):145–50.
31. Chipchase SY, Chapman HR, Bretherton R. A study to explore if dentists' anxiety affects their clinical decision-making. Br Dent J [Internet]. 2017;222(4):277–90. doi: 10.1038/sj.bdj.2017.173.
32. Plessas A, Nasser M, Hanoch Y, O'Brien T, Bernardes Delgado M, Moles D. Impact of time pressure on dentists' diagnostic performance. J Dent [Internet] 2019;82(January):38–44. doi: 10.1016/j.jdent.2019.01.011.
33. Plessas A, Delgado MB, Nasser M, Hanoch Y, Moles DR. Impact of stress on dentists' clinical performance. A systematic review. Community Dent Health 2018;35(1):9–15.
34. Junior O, Andrade FDE. Effect of clinical experience on accuracy and reliability of radiographic caries detection. Int J Odontostomatol 2017;11(3):347–52.
35. Hegde S, Gao J, Vasa R, Cox S. Factors affecting interpretation of dental radiographs. Dentomaxillofacial Radiol [Internet]. 2022;52(2):20220279. doi: 10.1259/dmfr.20220279.
36. Glaser BG, Strauss AL. Discovery of grounded theory: strategies for qualitative research [Internet]. New York: Taylor & Francis; 2017. Available from: https://books.google.com.au/books?id=GTMrDwAAQBAJ. Accessed 31 June 2022.
37. Waite S, Grigorian A, Alexander RG, et al. Analysis of perceptual expertise in radiology – current knowledge and a new perspective. Front Hum Neurosci [Internet] 2019;13:213. doi: 10.3389/fnhum.2019.00213.
38. Eli I, Weiss EI, Tzohar A, Littner MM, Gelernter I, Kaffe I. Interpretation of bitewing radiographs. Part 1 evaluation of the presence of approximal lesions. J Dent 1996;24(6):379–83.
39. Secgin CK, Gulsahi A, Arhun N. Diagnostic challenge: instances mimicking a proximal carious lesion detected by bitewing radiography. Oral Health Dent Manag 2016;15(6):2–6.
40. Baelum V. What is an appropriate caries diagnosis? Acta Odontol Scand 2010;68(2):65–79.
41. Banerji S, Mehta SB, Millar BJ. Cracked tooth syndrome. Part 1: aetiology and diagnosis. Br Dent J 2010;208(10):459–63.
42. Shackleton TA. Failure of root canal treatment misdiagnosed as neuropathic pain: case report. J Can Dent Assoc 2013;79:d94.
43. Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. Ulster Med J 2012;81(1):3–9.
44. Feller L, Lemmer J, Nemutandani MS, Ballyram R, Khammissa RAG. Judgment and decision-making in clinical dentistry. J Int Med Res [Internet] 2020;48(11):0300060520972877. doi: 10.1177/0300060520972877.
45. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. Am J Roentgenol 2013;201(3):611–7.
46. Krombach JW, Edwards WA, Marks JD, Radke OC. Checklists and other cognitive aids for emergency and routine anesthesia care-a survey on the perception of anesthesia providers from a large academic US institution. Anesthesiol pain Med 2015;5(4):e26300.
47. Weiss MJ, Kramer C, Tremblay S, Côté L. Attitudes of pediatric intensive care unit physicians towards the use of cognitive aids: a qualitative study. BMC Med Inform Decis Mak [Internet] 2016;16(1):53. doi: 10.1186/s12911-016-0291-6.
48. Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit*. Crit Care Med [Internet] 2004;32(10):2014–20.
49. Devlin H, Williams T, Graham J, Ashley M. The ADEPT study: a comparative study of dentists' ability to detect enamel-only proximal caries in bitewing radiographs with and without the use of AssistDent arti cial intelligence software. Br Dent J 2021;231(8):481–5.
50. Hajjaj FM, Salek MS, Basra MKA, Finlay AY. Non-clinical in uences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010;103(5):178–87.
51. Bos-Touwen ID, Trappenburg JCA, van der Wulp I, Schuurmans MJ, de Wit NJ. Patient factors that in uence clinicians' decision making in self-management support: a clinical vignette study. PLoS One 2017;12(2):e0171251.
52. George A, Ajwani S, Bhole S, et al. Knowledge, attitude and practises of dentists towards oral health care during pregnancy: a cross-sectional survey in New South Wales, Australia. Aust Dent J [Internet] 2017;62(3):301–10. doi: 10.1111/adj.12505.
53. Jadav NM, Abbott PV. Dentists' knowledge of dental trauma based on the International Association of Dental Traumatology guidelines: an Australian survey. Dent Traumatol 2022;38(5):374–80.
54. Bulmer JA, Currell SD, Peters CI, Peters OA. Endodontic knowledge, attitudes and referral patterns in Australian general dentists. Aust Dent J [Internet] 2022;67(S1):S24–30. doi: 10.1111/adj.12912.
55. Registrant Data 2023 [Internet]. Dental Board of Australia Registrant Data. 2023 [cited 2023 November 1]. Available from: https://www.dentalboard.gov.au/about-the-board/statistics.aspx

Table S1

Correlation among the various factors affecting interpretative errors and the frequency of interpretive errors. Correlation coefficient values are presented in this table.

		Clinician performance related factors			Patient-related factors		Image interpretation factors				Image quality-related factor
	Factors affecting errors of interpretation	Excessive workload	Dentist not alert	Dentist unwell	Complex case	Patients with dental fear	Under reading a radiograph	Jumping to conclusions	Satisfaction of search	Over-reliance on previous radiographs	Mistaking film fault for pathology
Demographic	Clinical experience				-0.258 *						
Clinician performance related factors	Excessive workload		0.562 **	0.594 **							
	Last patient of the day	0.488 **		0.571 **							
	Workplace distraction	0.517 **		0.515 **							
	Inexperienced dentist	0.446 **		0.426 **							
Patient related factors	Complex case					0.509 **					
	Demanding patient					0.583 **					
	When dental history is not obtained				0.418 **						
Image interpretation -related factors	Jumping to conclusions						0.588 **		0.498 **	**0.707** **	0.474 **
	Over-reliance on memory (seen on another radiograph)							0.557 **	0.655 **	0.582 **	
	Over-reliance on previous radiographic interpretation						0.582 **		0.587 **		0.451 **
	Satisfaction of search						0.514 **				
	Missing a lesion outside the area of interest						0.478 **	0.508 **	0.569 **		0.408 **
	Misdiagnosis due to lesion mimicking other lesions										0.415 **

* Indicates p- value <0.05; **indicates p-value<0.001

Supplementary 1: Survey Questionnaire

Dentist's Perceptions of Errors of Interpretation of Radiographs

Introduction and Survey information

Summary of the research project:

Errors in the interpretation of medical imaging are the most common causes leading to diagnostic error. In clinical dentistry, the rate and cause of error in interpretation of dental radiographs is not as well documented and studied. This survey aims to explore your perceptions of errors of interpretation on dental radiographs, causes and implications of these errors and strategies to minimise them. The participant information statement attached here describes the study in further detail.

This survey will take approximately 20 minutes to complete.

Participant Information Statement

Thank you for taking the time to participate in the survey!

Section 1

How important are **radiographs** in treatment planning for a patient?

Not at all important	Slightly important	Moderately important	Very important	Extremely important
○	○	○	○	○

How often do you encounter errors of radiology interpretation in your practice?

When an error of radiology interpretation occurs, is it documented?

- ◯ Yes
- ◯ No

When an error in radiology interpretation occurs, is the patient informed?

- ◯ Yes
- ◯ No

To what extent do you feel that errors of radiology interpretation are preventable?

Never	Sometimes	About half the time	Most of the time	Always
◯	◯	◯	◯	◯

Section 2: Frequency of errors

Please **rate in the order of likelihood**, type of error likely to occur in your practice?

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewha likely
1. **Error of Omission**: when no diagnosis is made	◯	◯	◯	◯
2. **Misdiagnosis**: when another diagnosis is made before the correct one	◯	◯	◯	◯
3. **Delayed Diagnosis**: when there is an unintentional delay in diagnosis	◯	◯	◯	◯
4. **Near Misses**: although an incident occurred that could have potentially lead to an adverse event if no harm came to the patient	◯	◯	◯	◯

Please rank the **conditions** where errors of interpretation on radiographs are likely to occur;

dental caries assessment of depth

cracked tooth misdiagnosed as sinusitis/TMD/atypical facial pain

misdiagnosis of late tooth development as hypodontia

failure to detect radiographic signs of periodontal disease

failure to detect radiographic signs of an odontogenic cyst

failure to detect radiographic signs of benign odontogenic tumour

cervical external root resorption misdiagnosed as root caries

failure to diagnose internal root resorption

any other cause :

Section 4: Contributing factors

Please rate the following **potential causes** for errors of radiology interpretation:

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely
1. When the dentist is not alert	○	○	○	○
2. When the dentist has an excessive workload	○	○	○	○
3. When treating the last patient of the day	○	○	○	○
4. When the dentist is unwell	○	○	○	○

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely
5. When the dentist is inexperienced- inadequate knowledge and skill	○	○	○	○
6. When a detailed history is not obtained	○	○	○	○
7. When treating a complex case	○	○	○	○
8. When treating a patient with dental fear	○	○	○	○
9. When treating a demanding patient	○	○	○	○
10. When treating a patient in pain	○	○	○	○
11. When there are distractions in the workplace	○	○	○	○

In your opinion, **how likely are the following to cause** errors of interpretation of radiographs ?

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewha likely
1. Reading a poor quality image	○	○	○	○
2. Mistaking a film fault for a pathology	○	○	○	○
3. Jumping to conclusions (not taking adequate time to analyse the radiograph)	○	○	○	○
4. Under-reading a radiograph(completely missing a lesion as a result)	○	○	○	○
5.Over reliance on previous interpretation of images	○	○	○	○
6. Over reliance on memory-may have seen a similar radiograph in the past	○	○	○	○
7. Failure to continue to analyse the image after an initial abnormality is detected	○	○	○	○
8. Completely missing lesions outside the area of interest	○	○	○	○

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewha likely
9. Misdiagnosis due to lesion mimicking another lesion	○	○	○	○
10. Underestimating the extent and severity of the lesion	○	○	○	○
11. Missing a diagnosis because of an inconspicuous lesion	○	○	○	○

Section 5: costs and consequences of errors of interpretations

In terms of **consequences of errors of interpretation to the patient**, please rate the likelihood of occurrence of following

	Extremely unlikely	Somewhat unlikely	Neither likely nor unlikely	Somewhat likely	Extremely likely
1. Severe harm or mortality	○	○	○	○	○
2. Over treatment	○	○	○	○	○
3. Under treatment	○	○	○	○	○
4. Negatively affecting patient's quality of life-physical, mental	○	○	○	○	○
5. Additional financial cost to the patient	○	○	○	○	○

In terms of **consequences of errors of interpretation to the dentist**, please rate the likelihood of occurrence of following

	Not at all important	Slightly important	Moderately important	Very important	Extremely important

	Not at all important	Slightly important	Moderately important	Very important	Extremely important
1. Additional cost to the dentist to retreat the patient	○	○	○	○	○
2. Loss of business/client	○	○	○	○	○
3. Loss of reputation	○	○	○	○	○
4. Legal implications	○	○	○	○	○
5. Increase in the risk profile-insurance premiums	○	○	○	○	○

SECTION 5: Strategies to reduce errors in interpretation of radiographs

Please **rate the following activities in the order of their significance in reducing errors of interpretation** in the context of each patient

	Not at all important	Slightly important	Moderately important	Very important	Extreme importar
1. Seeking and analysing a patient's previous radiographs	○	○	○	○	○
2. Comparing a patient's current(latest) radiograph to their previous radiographs	○	○	○	○	○
3. Ensuring high quality images are used for diagnosis	○	○	○	○	○
4. Prescribing appropriate radiographs based on the patient's complaint	○	○	○	○	○

Please **rate the significance** of the following in **reducing errors of interpretation** of dental radiographs

	Not at all important	Slightly important	Moderately important	Very importa
1. Using a checklist or a template for analysis of radiographic images to standardise the process	○	○	○	○
2. Having reflective practice to critically analyse clinical decisions	○	○	○	○
3. Phone consultation with an expert	○	○	○	○
4. Machine learning systems that detect errors of interpretation	○	○	○	○
5. Machine-learning systems to facilitate diagnostic feedback-automated decision support	○	○	○	○
6. Further education and training to enhance diagnostic skills	○	○	○	○
7. Discussion of cases with colleagues	○	○	○	○

SECTION 6 : Participant Demographics

Age

Gender

Type of clinical practice : select all that is applicable

- ☐ general
- ☐ public
- ☐ metropolitan
- ☐ none
- ☐ specialist
- ☐ private
- ☐ rural
- ☐

other

Please indicate **number of hours spent per week** doing the following activities

- [ ] clinical
- [ ] research
- [ ] teaching
- [ ] administration

Please indicate the number of **years of clinical experience**?

Please indicate average number of your patients requiring radiographs per day?

Please select the types of radiographs frequently taken in your practice **and** indicate the number of each radiograph taken per day

- [ ] periapical radiographs
- [ ] bitewings
- [ ] panoramic radiographs(OPG)
- [ ] other

Type of imaging system used for intraoral radiography in your practice

Photo stimulable direct digital chemically other

phosphor (PSP) ☐ (CCD/CMOS) ☐ processed/wet films ☐ ☐

Phase 2 of this study involves focus group interviews to gain in depth understanding of the decision making processes involved in diagnoses on dental radiographs. It will also involve participation in a short radiographic diagnosis activity.
Would like to participate in the second phase of this study?

- ◯ Yes
- ◯ No
- ◯ ☐ If yes, please enter your email address

Would you like to receive future communication with the results of the study?

- ☐ Yes
- ☐ No
- ☐ ☐ If yes, please enter email address

Powered by Qualtrics

Chapter 4

Australian Dentist's Knowledge of Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures

Understanding how dental practitioners perceive the consequences of these errors and the measures they consider effective in mitigating them is essential in improving diagnostic accuracy and patient care. This chapter presents a survey that assessed the perceptions regarding consequences and potential methods to minimise errors.

WILEY

Clinical and Experimental Dental Research
Open Access

ORIGINAL ARTICLE OPEN ACCESS

Australian Dentists' Knowledge of the Consequences of Interpretive Errors in Dental Radiographs and Potential Mitigation Measures

Shwetha Hegde[1] | Shanika Nanayakkara[2] | Stephen Cox[3] | Rajesh Vasa[4] | Jinlong Gao[2]

[1]Dentomaxillofacial Radiology, Sydney Dental School, University of Sydney, Sydney, Australia | [2]Sydney Dental School, Institute of Dental Research, Westmead Centre for Oral Health, University of Sydney, Sydney, Australia | [3]Discipline of Oral Surgery, Sydney Dental School, University of Sydney, Sydney, Australia | [4]Translational Research and Development, Applied Artificial Intelligence, Deakin University, Melbourne, Australia

Correspondence: Shwetha Hegde (shwetha.hegde@sydney.edu.au)

Received: 1 February 2024 | **Revised:** 9 August 2024 | **Accepted:** 1 October 2024

Funding: This study was supported by Australian Dental Research Foundation, Grant/Award Numbers: 2790, 2020.

Keywords: dental radiography | medical errors | patient safety | radiographic image interpretation

ABSTRACT

Objectives: Dental radiographs, typically taken and interpreted by dentists, are essential for diagnosis and effective treatment planning. Interpretive errors in dental radiographs, stemming from failures of visual and cognitive processes, can affect both patients and clinicians. This survey aimed to assess the dental practitioners' perceptions of the consequences of these errors and potential measures to minimize them.

Materials and Methods: This online anonymized survey assessed Australian dental practitioners' perceptions of the consequences of these errors and potential mitigation measures using ranking, Likert scale, and open-ended questions. The data were analyzed using descriptive statistics and bivariate analysis.

Results: Participants identified undertreatment (72%) and legal implications (82%) as the most significant consequences of interpretive errors, whereas severe harm to patients was deemed the least likely. Dental practitioners placed a greater emphasis on maintaining a high level of competence and the well-being of their patients. Utilizing high-quality images (63.9%) and appropriate radiographs (59.7%) were identified as the most effective measures to minimize interpretive errors. Participants showed hesitancy regarding the reliance on machine learning as a clinical decision-making tool.

Conclusions: The survey provides valuable practical insights into the consequences and targeted measures to minimize the occurrence of interpretive errors. Efforts to minimize interpretive errors should address patient safety and practitioners' concerns about professional reputation and business viability. The study also suggests further research into the role of machine learning algorithms in reducing interpretive errors in dentistry.

1 | Introduction

A great deal of attention and effort is dedicated to ensuring the provision of safe and high-quality care to patients. Nevertheless, mistakes arise from radiographic interpretive errors, leading to unintentional patient harm caused by medical management rather than the patient's underlying condition (Kohn, Corrigan, and Donaldson 2000). Interpretive errors are a type of diagnostic error that has been extensively studied in medical radiology (Degnan et al. 2019; Itri et al. 2018; Waite et al. 2017). Although interpretive errors in dental radiographs have not been as widely studied as in medical radiology, evidence suggests that they occur in dental practice, and several factors influence their occurrence (Hegde et al. 2024). These factors

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Clinical and Experimental Dental Research* published by John Wiley & Sons Ltd.

include time pressure, stress, clinical experience, case complexity, and cognitive load (Bretherton, Chapman, and Chipchase 2016; Cave and Hutchison 2019; Hegde et al. 2022; Plessas et al. 2018, 2019).

The severity of consequences arising from interpretive errors varies and depends on factors such as the type of error, the stage of the diagnostic process, and the promptness of identification and correction (Balogh et al. 2015). These errors can lead to adverse outcomes that can be classified as patient-related or clinician-related consequences, potentially resulting in unnecessary treatments, malpractice lawsuits, damage to the clinician's reputation and credibility, increased financial burdens for both the clinician and patient, and a negative impact on the patient's quality of life (Nikdel et al. 2018). Among the consequences of interpretive errors in medical imaging, malpractice lawsuits have been particularly significant (Busardò et al. 2015). Although extreme consequences of interpretive errors in dentistry are unlikely, it is essential to acknowledge and mitigate these errors to ensure quality patient care and minimize associated consequences. Furthermore, the context of dental imaging, radiographic interpretation, and diagnosis is distinct in dentistry compared to medicine. General dentists are responsible for capturing and interpreting dental plain film radiographs, whereas in medicine, specialists perform most imaging and interpretation of medical radiographs. This distinction could lead to consequences of interpretive errors that are unique to dentistry. Given their role in initial radiographic screening, general dentists play a crucial role in the early detection and diagnosis of dental conditions (Rushton, Horner, and Worthington 2001; Zeichner, Ruttimann, and Webber 1987).

With advances in diagnostics and healthcare technologies, healthcare is poised to minimize errors and improve diagnostic accuracy and quality of patient care. However, there is no single solution to the problem of interpretive errors. In medical imaging, several approaches have been applied to improve accuracy in radiographic diagnosis. These include an emphasis on critical analysis and clinical decision-making in medical education (Morrissey and Heilbrun 2017), the use of simulation in teaching diagnostic methods (Solomon et al. 2021), diagnostic checklists for structured analysis of medical images (Ely, Graber, and Croskerry 2011), and improvements in technology such as automated clinical decision support (Oakden-Rayner 2019) and machine learning (ML) algorithms (Choy et al. 2018). Using technology can be advantageous in minimizing interpretive errors as it can enhance human performance. Peer review with continuous and critical reflection of physicians' performance using structured procedures has also been shown to improve diagnostic accuracy (Lee et al. 2013). Active and effective management of interpretive errors requires the development of a system for identifying and managing these errors. Awareness and knowledge of the frequency and the causes are prerequisites for identifying interpretive errors as and when they occur in clinical practice. These steps will enable the implementation of effective measures to prevent such errors from occurring in the future.

This survey aims to gather dental practitioners' views on the impact of interpretive errors on both themselves and their patients. This study also aims to assess dental practitioners' perceptions of the currently available measures to mitigate such errors in clinical practice.

2 | Materials and Methods

A cross-sectional, anonymized survey of dentists practising in New South Wales, Australia, was undertaken. This survey was conducted as a part of a broader study that assessed the perceptions of Australian dental practitioners about interpretive errors in dentistry (Hegde et al. 2024). This study was approved by the University of Sydney Human Research Ethics Committee (Approval Number 2020/336) and conducted in accordance with the Declaration of Helsinki (World Medical Association 2013). The target population included dental practitioners within the research team's network, subscribers to the ADA's e-newsletters, users of social media platforms, and attendees of dental continuing professional development (CPD) workshops. An anonymized link to the survey was created and made available to potential participants via emails, e-newsletters, social media platforms, and during CPD workshops. Informed consent was assumed when the participants completed and submitted the survey.

The survey was conducted in two phases: the first from September 2020 to July 2021, and the second from January to March 2022, with the latter aimed at increasing participant numbers. Regular monitoring of the data during these periods allowed for the observation of response trends. Data saturation was determined manually, evidenced by the lack of new information arising from subsequent responses, a method analogous to the data saturation techniques employed in qualitative studies. The survey concluded once it was established that data saturation had been achieved, in line with the principles of grounded theory (Glaser and Strauss 2017). The online questionnaire was developed and distributed on the Qualtrics platform (www.qualtrics.com). After conducting an extensive review of literature pertaining to interpretive errors in radiology and clinical decision-making, as well as consulting experts in radiology and survey design, a questionnaire was created. This questionnaire aimed to explore how participants perceived the consequences of interpretive errors in dental radiographs and to explore their views on measures to minimize interpretive errors in clinical dentistry. The questionnaire comprised several categories, including questions about the consequences of interpretive errors for both clinicians and patients and measures to reduce such errors in dentistry. Demographic data gathered included participants' location and type of practice, the type and the number of radiographs taken in their dental practice each day, and the type of intra-oral imaging system used in their practice. The survey questionnaire is included in Supporting Information S1: File 1, with questions in Sections 5, 6, and 7 addressing this study. These sections included questions about the consequences of interpretive errors, potential solutions to minimize errors, and participant demographic data (Supporting Information Material).

Data were exported from the Qualtrics platform to SPSS statistical software (Version 26, IBM, SPSS, Chicago, IL) for analysis. Cronbach's α was used to measure the internal consistency

of the survey instrument. Descriptive statistics were used to summarize the demographic data of participants. Chi-squared tests of independence were used to find the association between categorical variables. Fisher's test was used when more than 20% of the cells had an expected frequency of less than 5%. Statistical comparisons were not conducted when expected cell counts were less than 1. The responses between different demographic groups, such as gender, type of practice (public or private), and level of training (general dentist or specialist), were compared. Spearman's rank correlation was used to test the correlations between variables. The statistical significance was set at $p < 0.05$.

3 | Results

This survey received 118 responses, of which 38 records were excluded from the analysis due to missing data. The survey instrument had a high internal consistency with Cronbach's α value of 0.928. The significant findings from the survey are presented here to highlight meaningful trends and insights.

3.1 | Demographic Features

A summary of the demographic characteristics of the survey participants is presented in Table 1. The participants' ages ranged from 23 to 72 years. The reported clinical experience ranged from 1 to 50 years, averaging 17.1 (±11.8) years. On average, the participants reported taking 12 radiographs a day: four periapical radiographs, five bitewings, and three OPGs. The participants reported taking a range of radiographs, and most of them used the photostimulable phosphor (PSP) imaging system.

3.2 | Consequences of Interpretive Errors Affecting Patients and Clinicians

Among the patient-related consequences, the participants recognized undertreatment (72%) as the most significant, followed by an increase in the financial burden to the patient (62%). In contrast, severe harm or patient mortality (30.7%) was considered the least likely outcome (Figure 1). For clinician-related consequences, participants identified legal consequences (82%) as the most likely ones. Loss of reputation (75.6%) was also recognized as a significant consequence, whereas the additional cost to the clinician (52.7%) was considered as the least likely consequence (Figure 2).

Statistically significant correlations were observed between patient-related and clinician-related consequences (Table S1). Additional cost to the dentist was the only one that correlated to all patient-related and clinician-related consequences with statistical significance ($p < 0.001$). Similarly, additional cost to the patient had a moderate correlation to all other patient-related consequences ($p < 0.001$). Interestingly, there was a strong positive correlation between the loss of a clinician's reputation and the loss of a client due to interpretive errors ($r = 0.770$, $p < 0.001$).

3.3 | Gender and Professional Qualification Variations in the Perceptions of the Consequences of Interpretive Errors

Interestingly, female and male participants had different opinions about the likelihood of patient-related and clinician-related consequences of interpretive errors. More female participants (78%) considered undertreatment as likely patient-related consequences of interpretive errors compared to male participants (63.6%). In contrast, more male participants considered patient harm or mortality (females, 40%; males, 30%) and negative impact on patient's quality of life (females, 40%; males, 28%) as the likely results of interpretive errors. When the clinician-related consequences were compared, more female participants (84.8%) considered that legal consequences were more likely consequences compared to the male participants (77.4%). Female participants also considered loss of client (78.4% for females and 40.9% for males, $p = 0.023$) and loss of reputation (females, 82.5%; males, 54.6%; $p = 0.039$) were more likely to occur than male participants.

When the responses of general dental practitioners (GP) were compared with the specialists, it was found that more GPs (74.3%) considered that undertreatment was the most likely patient-related outcome compared to specialist dentists (71.4%). More specialists (71%) thought that additional cost to the patient was a possible outcome compared to the GPs (60%). When the responses regarding the clinician-related consequences were compared, trends were revealed. More specialists (71%) considered a hike in insurance premiums was likely to occur due to interpretive errors compared to GPs (54%). In contrast, more GPs (65.7%) considered loss of client as a more likely consequence of interpretive error compared to specialists (57%). For each patient-related and clinician-related consequence, dentists practising in public practice (hospital-based) considered them more likely to occur than dentists in private practice.

3.4 | Measures to Minimize Errors

Participants identified that using high-quality images for diagnosis (63.9%) and prescribing appropriate radiographs based on the patient's chief complaint (59.7%) were the most effective means of minimizing interpretive errors. Undergoing training and education to enhance diagnostic skills (54.2%) and discussing a case with a colleague (45.8%) were also recognized as effective measures. Interestingly, although the use of ML algorithms was identified as the least effective measure (11.1%) for reducing the incidence of interpretive errors in dentistry (Figure 3), it was observed to correlate with measures related to cognitive aids such as standardized checklists ($p < 0.001$) (Table S2).

3.5 | Gender Perspectives Regarding Measures to Minimize Interpretive Errors

Although there were no significant gender differences in the responses, more male participants (72.7%) considered using high-quality radiographs as the essential measure to minimize interpretive errors compared to female participants (60.8%).

TABLE 1 | Demographic characteristics of the participants.

Characteristics	Value ($n = 80$)
Age (years)	
Mean (SD)	41.7 ± 11.0
Minimum	23
Maximum	72
Gender, *n* (%)	
Female	46 (57.5)
Male	22 (27.5)
Not specified	12 (15.0)
Clinical experience in years	
Mean (SD)	17.1 (±11.8)
Minimum	1
Maximum	50
Type of clinical practice (choice count) *n* (%)	
Type of practice[a]	
General	35 (43.7)
Specialist	14 (17.5)
Not specified	31 (38.8)
Sector[b]	
Private	21 (26.3)
Public	14 (17.5)
Both	11 (13.8)
Practice location[b]	
Metropolitan	13 (16.4)
Regional	5 (6.3)
Other (university academic)	2 (0.03)
Time spent doing different activities per week (mean hours [SD])	
Clinical	27.90 (±12.5)
Research	3.09 (±5.4)
Teaching	5.76 (±7.5)
Administration	4.83 (±5.9)
Patients per day requiring radiographs (mean ± SD)	6.2 (±2.5)
Number of radiographs taken per day (mean ± SD)	
Periapical	3.8 (±3.4)
Bitewings	4.9 (±3.7)
Panoramic radiographs (OPG)	2.7 (±2.7)
Other	0.49 (±1.0)
Imaging systems used in clinical practice, *n* (%) (choice count)[b]	
Photostimulable phosphor (PSP)	42 (52.5)
Direct digital (CCD, CMOS)	25 (31)
Analog films (chemically processed)	14 (17.5)
Other	20 (25)

[a]The total number of responses does not add up to 100% for some variables due to missing responses.
[b]As the participants have selected more than one choice, the responses don't add up to 100%.

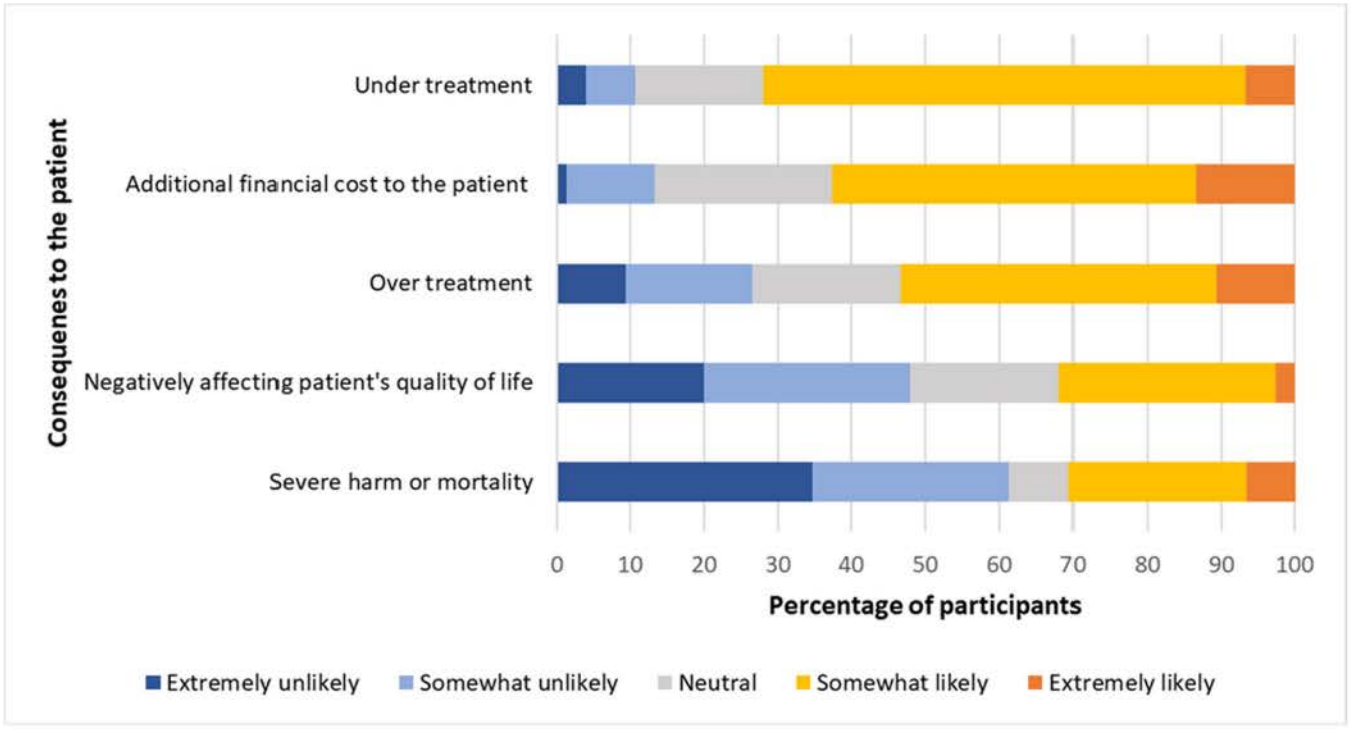


FIGURE 1 | Likelihood of consequences of interpretive errors affecting the patient.

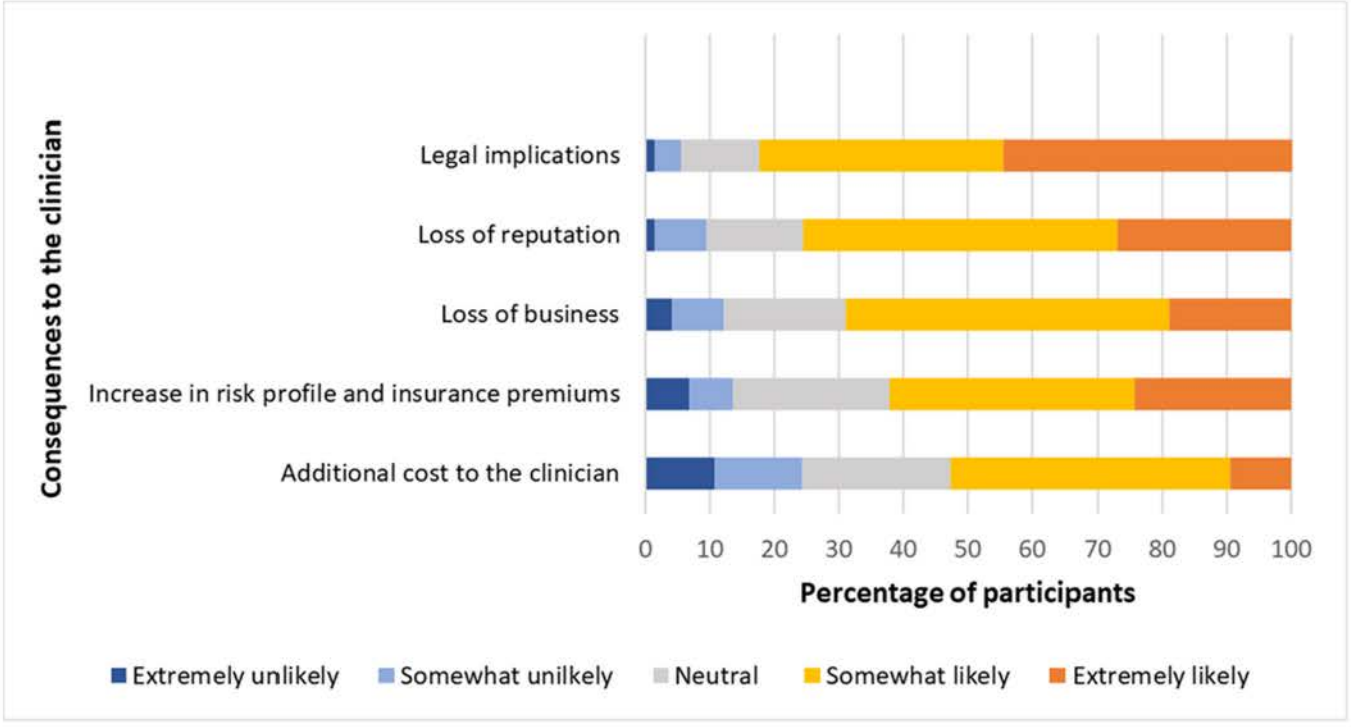


FIGURE 2 | Likelihood of consequences of interpretive errors affecting the clinician.

More male participants (68%) considered prescribing appropriate radiographs a more effective error-minimizing measure than females (58.6%). However, more female participants (43%) considered comparing new radiographs with old ones as an effective measure compared to males (36%). There were no differences in the responses between GPs and specialists and between dentists in public and private practices.

4 | Discussion

The survey aimed to assess the dental practitioners' perceptions about the impact of interpretive errors in dentistry and potential solutions to minimize the occurrence of such errors. One notable observation was that the participants generally did not perceive interpretive errors led to serious harm to the patient.

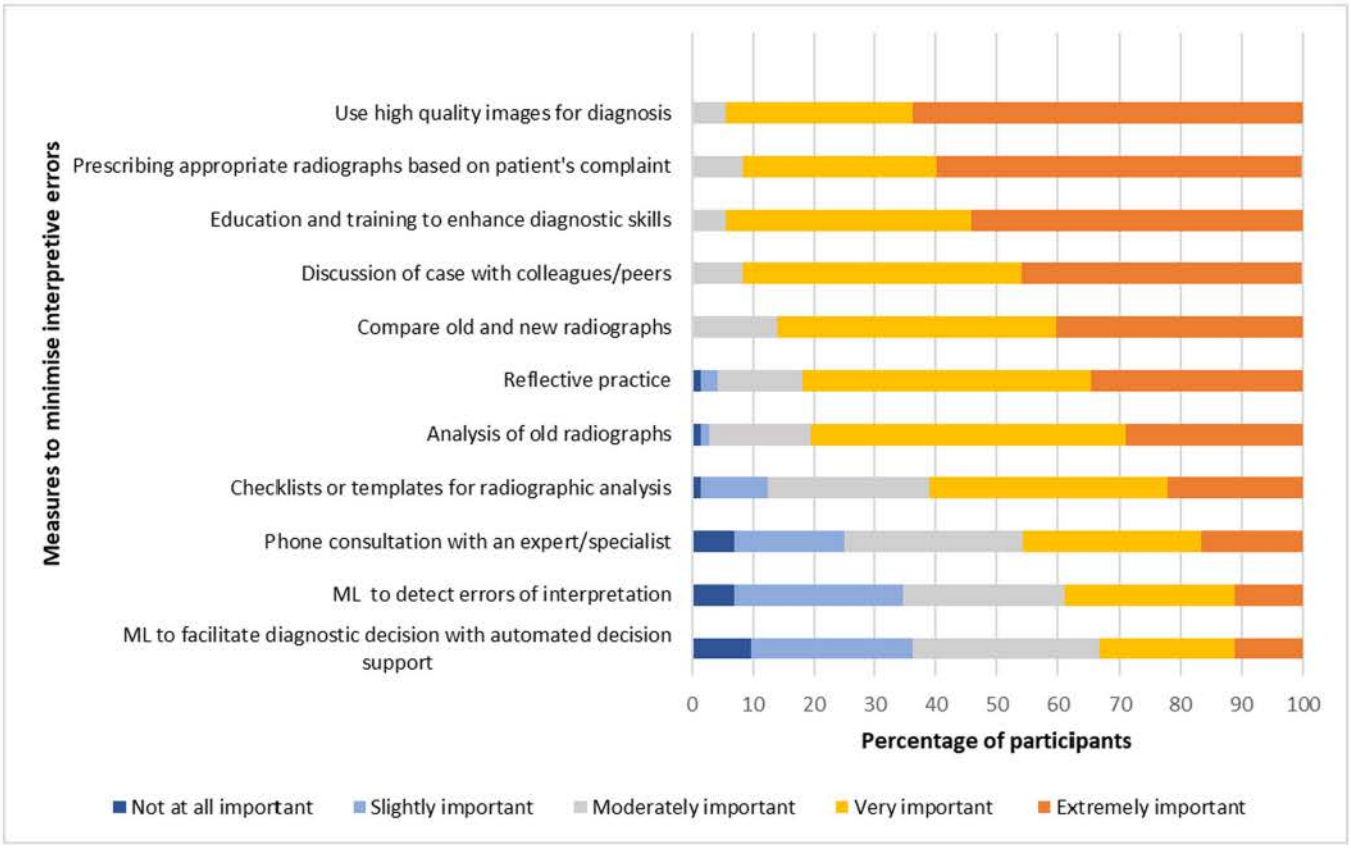


FIGURE 3 | Measures to minimize interpretive errors.

Instead, the survey revealed that dental practitioners placed greater emphasis on the potential consequences of interpretive errors in their practice, particularly in terms of the loss of reputation and financial consequences. This finding suggests that practitioners were more concerned about the impact of errors on their professional standing and the potential negative effects on their relationships with patients and colleagues.

The emphasis on loss of business and reputation as significant consequences of interpretive errors highlights the importance that dental practitioners place on maintaining a high level of competence and accuracy in their diagnostic and interpretive abilities. It underscores the potential impact that errors can have on patient trust and satisfaction and the potential for negative word-of-mouth and reputational damage. These findings contrast with medical radiology, where malpractice lawsuits are the biggest concern as a consequence of interpretive errors (Berlin 2017; Biswas et al. 2023). Other consequences of interpretive errors in medical radiology include repeat imaging, delayed or inappropriate treatment, avoidable hospital admissions, and increased mortality (Ahn et al. 2021; Onder et al. 2021). Overall, dental and medical practitioners share a common goal of minimizing interpretive errors. However, the concerns of dental practitioners are distinct from those of medical practitioners, which could be attributed to the unique contexts and priorities within each healthcare domain, such as variations in patient expectations, scope and impact of treatment provided, clinical practice dynamics, and legal considerations.

In countries like Australia, dental radiology diagnosis is mainly performed by general dentists. Admittedly, the findings of this study may not be generalizable to those countries with a higher number of DMFR specialists who provide radiological diagnosis and reporting. However, it is important to note that because general dentists perform routine dental images such as intraoral and panoramic radiographs and CBCT images, they are in a crucial position to detect jaw pathologies and malignancies at an early stage (MacDonald and Yu 2020; Price et al. 2012). Early detection of these conditions is essential to prevent disease progression, extensive destruction of dental supporting structures, and potential patient morbidity (Choi 2011). Therefore, targeted training and improving the diagnostic skills of general dentists can help minimize missed lesions, particularly because early-stage diseases are often asymptomatic and may be identified incidentally on dental radiographs.

The survey also provided valuable insights into the interrelationships among the various consequences of interpretive errors. Participants considered that overtreatment negatively impacted their patients' quality of life and resulted in additional financial burdens on the clinician, legal consequences, and loss of the clinician's reputation. This highlighted the participant's awareness that negative patient outcomes were linked to financial and legal consequences for the clinicians.

The participants' perceptions of measures to minimize interpretive errors were also analyzed in this survey. The findings emphasize the significance of utilizing high-quality images (taken using the correct technique and exposure settings and

having optimal contrast, density, and brightness) and choosing appropriate radiographs based on patient symptoms. The impact of image quality has been investigated in medicine and dentistry (Boita et al. 2021; Bruno, Walker, and Abujudeh 2015; Ghazali, Mohd Yusof, and Norman 2021). These practices are crucial for improving diagnostic accuracy and reducing interpretive errors in dentistry. However, it is essential to consider that these higher quality images may not always align with the "As Low As Reasonably Achievable" (ALARA) principle, which prioritizes minimizing patient exposure. Further studies are needed to develop imaging protocols for specific dental conditions that balance the diagnostic efficacy with minimum radiation dose. Studies have shown that there is a greater likelihood of interpretive errors with 3D imaging compared to 2D imaging (Williams and Drew 2019). However, this survey did not study the impact that 2D versus 3D imaging may have on the incidence and consequences of interpretive errors.

Among the different measures to minimize interpretive errors, the participants also recognized training, education, and consulting with peers or colleagues as the most effective strategies. This emphasizes the importance of CPD and recognizes the value of collaboration and seeking input from others to enhance the accuracy of diagnosis and mitigate individual biases or oversights. Previous studies have shown that CPD can enhance clinical decision-making skills and patient outcomes (Firmstone et al. 2013). The survey results could inform the design of training programs and implement quality assurance measures to minimize interpretive errors, improve patient care, and safeguard the professional reputation of dental clinicians.

Interestingly, the use of ML algorithms as cognitive aid was rated as the least effective measure for reducing the occurrence of interpretive errors, potentially due to scepticism about ML in dentistry. Studies have suggested that this scepticism may arise from concerns such as the interpretability and explainability of ML algorithms (Antoniadi et al. 2021). In addition, there are concerns that automated algorithms cannot fully replace or replicate the nuanced decision-making capabilities of experienced dental professionals (Ding et al. 2023). There may also be concerns about patient privacy (Pethani 2021), liability, and responsibility for decisions made by the automated ML algorithms (Naik et al. 2022). It is important to note that this survey was conducted before the wider applications of ML algorithms and the development of large language models using generative AI. Given the advancements in technology since then, it is reasonable to consider that the perceptions of ML algorithms may have evolved if the survey were conducted in the present time.

The study's limitations include a small sample size, which may preclude the generalizability of the survey results. However, similar low participant numbers were noted in other Australian surveys of dental practitioners (Bulmer et al. 2022; George et al. 2017; Jadav and Abbott 2022). In addition, the low GP to dental specialist ratio may have limited the interpretation of the results in this survey, especially when comparisons were made between these two groups. Although the survey responses provided valuable insights, not all questions were answered by every participant. Future research will aim to validate the suggestions provided by the participants and explore the correlation between participant's perceptions about interpretive errors and their age and clinical experience. This correlation was not examined in this study due to the low sample size, which resulted in fewer participants in each age and clinical experience category. Furthermore, as this survey was voluntary, this may have resulted in selection bias.

5 | Conclusions

The survey findings highlight the perceptions and beliefs of dental practitioners regarding the consequences of interpretive errors in dental radiology and potential solutions to minimize their occurrence. The results provide valuable insights to develop targeted interventions based on the clinician's level of training, gender, and sector of practice. Efforts to minimize interpretive errors should not only focus on improving patient safety but also address the concerns and priorities of dental practitioners, including the preservation of their professional reputation and business viability. However, there remains scepticism about the effectiveness of ML algorithms in dentistry. Future research is needed to explore the evolving perceptions of ML algorithms in dentistry and address these concerns effectively.

Author Contributions

Shwetha Hegde has made substantial contributions to conceptualization, design, project administration, investigation and formal analysis, visualization, writing the original draft, and revising the manuscript. Jinlong Gao is involved in methodology, feedback on the concept, data analysis and writing (review and editing), and visualization. Stephen Cox participated in methodology, feedback on the concept, and writing (review and editing). Shanika Nanayakkara contributed to data curation and formal analysis, visualization, and writing (review and editing). Rajesh Vasa contributed to methodology, feedback on concepts, and writing (review and editing).

Acknowledgments

This project was partly supported by the Australian Dental Research Fund (grant number 2790, 2020). Open access publishing facilitated by The University of Sydney as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Ahn, Y., G.-S. Hong, K. J. Park, C. W. Lee, J. H. Lee, and S.-O. Kim. 2021. "Impact of Diagnostic Errors on Adverse Outcomes: Learning From Emergency Department Revisits With Repeat CT or MRI." *Insights Into Imaging* 12, no. 1: 160. https://doi.org/10.1186/s13244-021-01108-0.

Antoniadi, A. M., Y. Du, Y. Guendouz, et al. 2021. "Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical

Decision Support Systems: A Systematic Review." *Applied Sciences* 11, no. 11: 5088. https://doi.org/10.3390/app11115088.

Balogh, E. P., B. T. Miller, J. R. Ball, and The National Academies of Sciences, Engineering, and Medicine. 2015. "Improving Diagnosis in Health Care." In *Improving Diagnosis in Health Care*, edited by E. P. Balogh, B. T. Miller, and J. R. Ball. Washington, DC: The National Academies Press. https://doi.org/10.17226/21794.

Berlin, L. 2017. "Medical Errors, Malpractice, and Defensive Medicine: An Ill-Fated Triad." *Diagnosis* 4, no. 3: 133–139. https://doi.org/10.1515/dx-2017-0007.

Biswas, S., S. Biswas, S. S. Awal, and H. Goyal. 2023. "Malpractice and Lawsuits in Radiology: What the Radiologist Should Know." *Egyptian Journal of Radiology and Nuclear Medicine* 54, no. 1: 19. https://doi.org/10.1186/s43055-023-00971-9.

Boita, J., R. E. van Engen, A. Mackenzie, et al. 2021. "How Does Image Quality Affect Radiologists' Perceived Ability for Image Interpretation and Lesion Detection in Digital Mammography?" *European Radiology* 31, no. 7: 5335–5343. https://doi.org/10.1007/s00330-020-07679-8.

Bretherton, R., H. R. Chapman, and S. Chipchase. 2016. "A Study to Explore Specific Stressors and Coping Strategies in Primary Dental Care Practice." *British Dental Journal* 220, no. 9: 471–478. https://doi.org/10.1038/sj.bdj.2016.334.

Bruno, M. A., E. A. Walker, and H. H. Abujudeh. 2015. "Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction." *Radiographics* 35, no. 6: 1668–1676. https://doi.org/10.1148/rg.2015150023.

Bulmer, J., S. Currell, C. Peters, and O. Peters. 2022. "Endodontic Knowledge, Attitudes and Referral Patterns in Australian General Dentists." Supplement, *Australian Dental Journal* 67, no. S1: S24–S30. https://doi.org/10.1111/adj.12912.

Busardò, F. P., P. Frati, A. Santurro, S. Zaami, and V. Fineschi. 2015. "Errors and Malpractice Lawsuits in Radiology: What the Radiologist Needs to Know." *La Radiologia Medica* 120, no. 9: 779–784. https://doi.org/10.1007/s11547-015-0561-x.

Cave, V., and C. Hutchison. 2019. "Does Time Pressure Impact on Dentists' Diagnostic Performance?" *Evidence-Based Dentistry* 20, no. 3: 81–82. https://doi.org/10.1038/s41432-019-0043-4.

Choi, J.-W. 2011. "Assessment of Panoramic Radiography as a National Oral Examination Tool: Review of the Literature." *Imaging Science in Dentistry* 41, no. 1: 1–6. https://doi.org/10.5624/isd.2011.41.1.1.

Choy, G., O. Khalilzadeh, M. Michalski, et al. 2018. "Current Applications and Future Impact of Machine Learning in Radiology." *Radiology* 288, no. 2: 318–328. https://doi.org/10.1148/radiol.2018171820.

Degnan, A. J., E. H. Ghobadi, P. Hardy, et al. 2019. "Perceptual and Interpretive Error in Diagnostic Radiology—Causes and Potential Solutions." *Academic Radiology* 26, no. 6: 833–845. https://doi.org/10.1016/j.acra.2018.11.006.

Ding, H., J. Wu, W. Zhao, J. P. Matinlinna, M. F. Burrow, and J. K. H. Tsoi. 2023. "Artificial Intelligence in Dentistry—A Review." *Frontiers in Dental Medicine* 4: 1–13. https://doi.org/10.3389/fdmed.2023.1085251.

Ely, J. W., M. L. Graber, and P. Croskerry. 2011. "Checklists to Reduce Diagnostic Errors." *Academic Medicine* 86, no. 3: 307–313. https://doi.org/10.1097/ACM.0b013e31820824cd.

Firmstone, V. R., K. M. Elley, M. T. Skrybant, A. Fry-Smith, S. Bayliss, and C. J. Torgerson. 2013. "Systematic Review of the Effectiveness of Continuing Dental Professional Development on Learning, Behavior, or Patient Outcomes." *Journal of Dental Education* 77, no. 3: 300–315. https://doi.org/10.1002/j.0022-0337.2013.77.3.tb05471.x.

George, A., S. Ajwani, S. Bhole, et al. 2017. "Knowledge, Attitude and Practises of Dentists Towards Oral Health Care During Pregnancy: A Cross Sectional Survey in New South Wales, Australia." *Australian Dental Journal* 62, no. 3: 301–310. https://doi.org/10.1111/adj.12505.

Ghazali, L., M. Y. P. Mohd Yusof, and N. H. Norman. 2021. "Effects of Scanning Parameters Reduction in Dental Radiographs on Image Quality and Diagnostic Performance: A Randomised Controlled Trial." *Journal of Orthodontics* 48, no. 1: 5–12. https://doi.org/10.1177/1465312520971641.

Glaser, B. G., and A. L. Strauss. 2017. *Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Routledge. https://books.google.com.au/books?id=GTMrDwAAQBAJ.

Hegde, S., J. Gao, R. Vasa, and S. Cox. 2022. "Factors Affecting Interpretation of Dental Radiographs." *Dentomaxillofacial Radiology* 52, no. 2: 20220279. https://doi.org/10.1259/dmfr.20220279.

Hegde, S., J. Gao, R. Vasa, S. Nanayakkara, and S. Cox. 2024. "Australian Dentist's Knowledge and Perceptions of Factors Affecting Radiographic Interpretation." *International Dental Journal* 74: 589–596. https://doi.org/10.1016/j.identj.2023.11.006.

Itri, J. N., R. R. Tappouni, R. O. McEachern, A. J. Pesch, and S. H. Patel. 2018. "Fundamentals of Diagnostic Error in Imaging." *RadioGraphics* 38, no. 6: 1845–1865. https://doi.org/10.1148/rg.2018180021.

Jadav, N. M., and P. V. Abbott. 2022. "Dentists' Knowledge of Dental Trauma Based on the International Association of Dental Traumatology Guidelines: An Australian Survey." *Dental Traumatology* 38, no. 5: 374–380. https://doi.org/10.1111/edt.12761.

Kohn, L. T., J. M. Corrigan, and M. Donaldson. 2000. "To Err is Human: Building a Safer Health System." In *Institute of Medicine (US) Committee on Quality of Health Care in America*. US: National Academies Press. https://doi.org/10.1016/j.yrtph.2007.09.017.

Lee, C. S., P. G. Nagy, S. J. Weaver, and D. E. Newman-Toker. 2013. "Cognitive and System Factors Contributing to Diagnostic Errors in Radiology." *American Journal of Roentgenology* 201, no. 3: 611–617. https://doi.org/10.2214/AJR.12.10375.

MacDonald, D., and W. Yu. 2020. "Incidental Findings in a Consecutive Series of Digital Panoramic Radiographs." *Imaging Science in Dentistry* 50, no. 1: 53–64. https://doi.org/10.5624/isd.2020.50.1.53.

Morrissey, B., and M. E. Heilbrun. 2017. "Teaching Critical Thinking in Graduate Medical Education: Lessons Learned in Diagnostic Radiology." *Journal of Medical Education and Curricular Development* 4: 238212051769649. https://doi.org/10.1177/2382120517696498.

Naik, N., B. M. Z. Hameed, D. K. Shetty, et al. 2022. "Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?" *Frontiers in Surgery* 9: 862322. https://doi.org/10.3389/fsurg.2022.862322.

Nikdel, C., K. Nikdel, A. Ibarra-Noriega, E. Kalenderian, and M. F. Walji. 2018. "Clinical Dental Faculty Members' Perceptions of Diagnostic Errors and How to Avoid Them." *Journal of Dental Education* 82, no. 4: 340–348. https://doi.org/10.21815/JDE.018.037.

Oakden-Rayner, L. 2019. "The Rebirth of CAD: How Is Modern AI Different From the CAD We Know?" *Radiology: Artificial Intelligence* 1, no. 3: e180089. https://doi.org/10.1148/ryai.2019180089.

Onder, O., Y. Yarasir, A. Azizova, G. Durhan, M. R. Onur, and O. M. Ariyurek. 2021. "Errors, Discrepancies and Underlying Bias in Radiology With Case Examples: A Pictorial Review." *Insights Into Imaging* 12, no. 1: 51. https://doi.org/10.1186/s13244-021-00986-8.

Pethani, F. 2021. "Promises and Perils of Artificial Intelligence in Dentistry." *Australian Dental Journal* 66, no. 2: 124–135. https://doi.org/10.1111/adj.12812.

Plessas, A., M. B. Delgado, M. Nasser, Y. Hanoch, and D. R. Moles. 2018. "Impact of Stress on Dentists' Clinical Performance. A Systematic Review." *Community Dental Health* 35, no. 1: 9–15. https://doi.org/10.1922/CDH_4118Plessas07.

Plessas, A., M. Nasser, Y. Hanoch, T. O'Brien, M. Bernardes Delgado, and D. Moles. 2019. "Impact of Time Pressure on Dentists' Diagnostic Performance." *Journal of Dentistry* 82, no. January: 38–44. https://doi.org/10.1016/j.jdent.2019.01.011.

Price, J. B., K. L. Thaw, D. A. Tyndall, J. B. Ludlow, and R. J. Padilla. 2012. "Incidental Findings From Cone Beam Computed Tomography of the Maxillofacial Region: A Descriptive Retrospective Study." *Clinical Oral Implants Research* 23, no. 11: 1261–1268. https://doi.org/10.1111/j.1600-0501.2011.02299.x.

Rushton, V., K. Horner, and H. Worthington. 2001. "Screening Panoramic Radiology of Adults in General Dental Practice: Radiological Findings." *British Dental Journal* 190, no. 9: 495–501. https://doi.org/10.1038/sj.bdj.4801014.

Solomon, A. J., R. W. England, A. R. Kolarich, and R. P. Liddell. 2021. "Disrupting the Education Paradigm: An Opportunity to Advance Simulation Training in Radiology—Radiology in Training." *Radiology* 298, no. 2: 292–294. https://doi.org/10.1148/radiol.2020203534.

Waite, S., J. Scott, B. Gale, **T**. Fuchs, S. Kolla, and D. Reede. 2017. "Interpretive Error in Radiology." *American Journal of Roentgenology* 208, no. 4: 739–749. https://doi.org/10.2214/AJR.16.16963.

Williams, L. H., and T. Drew. 2019. "What Do We Know About Volumetric Medical Image Interpretation?: A Review of the Basic Science and Medical Image Perception Literatures." *Cognitive Research: Principles and Implications* 4, no. 1: 21. https://doi.org/10.1186/s41235-019-0171-6.

World Medical Association. 2013. "World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects." *JAMA* 310, no. 20: 2191–2194. https://doi.org/10.1001/jama.2013.281053.

Zeichner, S. J., U. E. Ruttimann, and R. L. Webber. 1987. "Dental Radiography: Efficacy in the Assessment of Intraosseous Lesions of the Face and Jaws in Asymptomatic Patients." *Radiology* 162, no. 3: 691–695. https://doi.org/10.1148/radiology.162.3.3809483.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

	Technological proficiency	Are you currently aware of any AI software used in dentistry?	How far in the future do you think it will be before AI is integrated into dentistry?	Do you think that applications of AI in dentistry should be a part of the dental and oral health curriculum?	How do you see AI impacting your future workflow, tasks and/or practice of dentistry and oral health? AI will;	How do you think artificial intelligence will change the role of the clinician? AI will;	If AI were integrated into your clinic within the year, the likelihood of you using this technology is;	The idea of AI being integrated into your practice of dentistry and oral health;	At what level of proficiency do you expect the AI to perform?	If your clinical judgement and AI differ, would you; - Selected Choice
Participant's age		-0.287**	-0.231**							
Technological proficiency					-0.176*					
Are you currently aware of any AI software used in dentistry?			0.384**	0.273**	0.178*	0.175*	**0.418****	0.254*		
How far in the future do you think it will be before AI is integrated into dentistry?					0.278**	0.248**	0.280*	0.245*		
How do you see AI impacting your future workflow, tasks and practice of dentistry?						**0.425****	**0.530****	0.387**	-0.163*	
How do you think AI will change the role of the clinician?							0.387**	0.397**	-0.223**	
If AI were integrated into your clinic within the year, the likelihood of you using this technology is								**0.453****		
At what level of proficiency do										0.164*

you expect the AI to perform?										

Table S1: Correlation between various aspects of perception of dental AI technologies

* Indicates p- value <0.05; **indicates p-value<0.001

Supplementary File 1

Survey Flow

Block: Default Question Block (33 Questions)

Page Break

Start of Block: Default Question Block

Q1
Thank you for your interest in this study.

Please read the Participant Information Statement attached before continuing. By submitting your survey, you consent to take part in the study, and you are consenting to the use of your information. Once responses are submitted, they cannot be withdrawn as surveys are completed anonymously.

Participant information statement

○ I have read the participant information statement and I consent to participating in this study (1)

Q2 **Attitudes and Perceptions of Australian Dental Practitioners and Dental Students to Applications of Artificial Intelligence in Dentistry: A Survey**
Artificial Intelligence (AI) refers to a system which attempts to simulate the human cognition and critical thinking, through interconnected computer processes which may have the ability to self-learn.

The application of artificial intelligence to dentistry is an emerging area of interest. Current research focuses on where and how this developing technology can be applied.

Due to its novel nature, there is not yet an established consensus on how the dental community will welcome this technology.This survey aims to investigate the attitudes of dental practitioners and students on this topic. **disclaimer; for the purposes of this study we use 'AI' to encompass the terms and concepts of machine learning, deep learning, artificial intelligence, and algorithms.*

Page Break

Q3 Please select one option below
I am a;

- ○ Dental Student (1)
- ○ Oral Health Student (2)
- ○ Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist) (3)
- ○ Non-Registered Australian Dental Practitioner (4)
- ○ Other (5)

Skip To: End of Survey If Please select one option below I am a; = Non-Registered Australian Dental Practitioner

Skip To: End of Survey If Please select one option below I am a; = Other

Q4 Have you heard/read of AI (Artificial Intelligence) in dentistry?

- ○ Yes (1)
- ○ No (2)

Skip To: Q6 If Have you heard/read of AI (Artificial Intelligence) in dentistry? = No

Q5 Please select and option below for where you heard/read about it
You may select more than one option

- [ ] In clinical practice (1)
- [ ] Social media (Facebook, Twitter, LinkedIn etc.) (2)
- [ ] Media (Newspaper, News, Podcast etc.) (3)
- [ ] Lectures/Education/Dental Journals/Conferences (4)
- [ ] Patients (5)
- [ ] Other - please specify (6) _____

Q6 Are you currently aware of any AI software used in dentistry?

- ○ Yes (1)
- ○ No (2)

Skip To: Q8 If Are you currently aware of any AI software used in dentistry? = No

Q7 Please select how you were made aware of this technology
You may select more than one option

- [ ] In clinical practice (1)
- [ ] Social media (Facebook, Twitter, LinkedIn etc.) (2)
- [ ] Media (Newspaper, News, Podcast etc.) (3)
- [ ] Lectures/Education/Dental Journals/Conferences (4)
- [ ] Patients (5)
- [ ] Other - please specify (6) _____

Q8 How far in the future do you think it will be before AI is integrated into dentistry?

- ○ Already Integrated (1)
- ○ Within a year (2)
- ○ 1-5 years (3)
- ○ 5-10 years (4)
- ○ Never (5)
- ○ Not sure (6)

Q9 Do you think that applications of AI in dentistry should be a part of the dental and oral health curriculum?

- ○ Yes (1)
- ○ No (2)
- ○ Not sure (3)

Skip To: Q11 If Do you think that applications of AI in dentistry should be a part of the dental and oral health... = No

Skip To: Q12 If Do you think that applications of AI in dentistry should be a part of the dental and oral health... = Not sure

Q10 Please specify why you think that AI should be incorporated into the dental and oral health curriculum?

_____

Skip To: Q12 If Condition: Please specify why you thin... Is Displayed. Skip To: In which field do you think AI will b....

Q11 Please specify why you think AI should not be incorporated into the dental and oral health curriculum

_____

Q12 In which field do you think AI will be most applicable? You may select multiple options

- [ ] Dental Maxillofacial Radiology (1)
- [ ] Implantology (2)
- [ ] Restorative Dentistry (3)
- [ ] Endodontics (4)
- [ ] Oral Surgery (5)
- [ ] Orthodontics (6)
- [ ] Periodontics (7)
- [ ] Prosthodontics (8)
- [ ] Oral Pathology and Oral Medicine (9)
- [ ] Special Needs Dentistry (10)
- [ ] Paediatric Dentistry (11)
- [ ] Other - please specify (12) _____

Q13 In what task do you think AI would be most beneficial? - please click and drag the options from most to least beneficial

_____ Administrative Tasks - booking patients, taking payments (1)
_____ Image Processing - tooth numbering, anatomical structure delineation (2)
_____ Diagnosis - Analysis and localisation of pathology and presentation of differential diagnoses (3)
_____ Prognosis - prognostic tool to predict the course of disease and/or success of treatment (4)
_____ Treatment planning tool - risk evaluations, positioning of implants (5)

Q14 How do you see AI impacting your future workflow, tasks and/or practice of dentistry and oral health?
AI will;

- ○ Benefit all of dentistry (1)
- ○ Benefit some aspects of dentistry (2)
- ○ Make no difference to dentistry (3)
- ○ Negatively impact some aspects of dentistry (4)
- ○ Negatively impact all of dentistry (5)

Q15 How do you think artificial intelligence will change the role of the clinician?

AI will;

- ○ Assist and support the clinician (1)
- ○ Somewhat assist the clinician (2)
- ○ Have no impact on the clinician (3)
- ○ Somewhat negatively affect the clinician (4)
- ○ Negatively affect the clinician (5)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q16 If AI were integrated into your clinic within the year, the likelihood of you using this technology is;

- ○ Highly Likely (1)
- ○ Somewhat Likely (2)
- ○ Unchanged (3)
- ○ Somewhat Unlikely (4)
- ○ Highly Unlikely (5)

Display This Question:

If Please select one option below I am a; = Oral Health Student

Or Please select one option below I am a; = Dental Student

Q17 If AI were integrated into your education curriculum within the year, the likelihood of you applying this technology is;

- ○ Highly Likely (1)
- ○ Somewhat Likely (2)
- ○ Unchanged (3)
- ○ Somewhat Unlikely (4)
- ○ Highly Unlikely (5)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q18 The idea of AI being integrated into your practice of dentistry and oral health;

- ○ Inspires excitement (1)
- ○ Makes little difference (2)
- ○ Raises concerns (3)

Skip To: Q21 If The idea of AI being integrated into your practice of dentistry and oral health; != Raises concerns

Display This Question:

If Please select one option below I am a; = Dental Student

Or Please select one option below I am a; = Oral Health Student

Q19 The idea of AI being integrated into your study of dentistry and oral health;

- ○ Inspires excitement (1)
- ○ Makes little difference (2)
- ○ Raises Concerns (3)

Skip To: Q21 If The idea of AI being integrated into your study of dentistry and oral health; != Raises Concerns

Q20 The main concerns you would have are around;
You can select multiple options

- [ ] Security - patient confidentiality, risk of hackers (1)
- [ ] Redundancy - practitioners losing jobs to more efficient technology (2)
- [ ] Legality/Insurance - who is liable (3)
- [ ] Trust - inadequate information/expertise (4)
- [ ] Patient care - inflexibility of programme, basis on physiology/pathology alone (5)
- [ ] Accuracy - chances of AI being incorrect/making mistakes (6)
- [ ] Other - please specify (7) _____

Q21 At what level of proficiency do you expect the AI to perform?

- ○ Equivalent or superior to the worst performing specialist (1)
- ○ Equivalent to the average performing specialist (2)
- ○ Superior to the average performing specialist (3)
- ○ Equivalent to the best performing specialist (4)
- ○ Superior to the best performing specialist (5)

Q22 If your clinical judgement and AI differ, would you;

- ○ Defer to the AI (1)
- ○ Trust your own clinical judgement (2)
- ○ Consult with a colleague or another experienced clinician (3)
- ○ Perform literature research (4)
- ○ Other - please specify (5) _____

Page Break

Q23
Demographics Please specify your age

_____

Q24 Please indicate your gender

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q25 Please select your dental field

- ○ General Practitioner (1)
- ○ Specialist (2)

Skip To: Q27 If Please select your dental field = General Practitioner

Skip To: Q29 If Please select your dental field = Specialist

Display This Question:

If Please select one option below I am a; = Dental Student

Or Please select one option below I am a; = Oral Health Student

Q26 Please select the type of degree and indicate which year of study you are currently in

- ○ Undergraduate - please indicate your current year of study (1)
  _____
- ○ Postgraduate - please indicate your current year of study (2)
  _____

Skip To: Q33 If Please select the type of degree and indicate which year of study you are currently in , Undergraduate - please indicate your current year of study Is Displayed

Q27 How many years have you been practicing?

_____

Q28 How many days do you work per week?

_____

Skip To: Q31 If Condition: How many days do you work p... Is Displayed . Skip To: Please select the options which descr....

Q29 Which field is your specialty?

_____

Q30 How many years have you been practicing?

_____

Q31 Please select the options which describe your current practice
You may select multiple options

- [ ] Metropolitan (1)
- [ ] Rural (2)
- [ ] Public (3)
- [ ] Private (4)
- [ ] Academia (5)

Q32 Please select all options which your current practice has in house
You may select multiple options

- [ ] OPG (1)
- [ ] CBCT (2)
- [ ] Digital Records (3)
- [ ] Digital Xray's (4)
- [ ] Cerec Machine (5)
- [ ] 3D scanner (6)

Q33 Please use your mouse to slide the bar to indicate how proficient you feel with technology in general

Unsatisfactory		Beginner		Satisfactory	Proficient
0	1	2	3	4	5

Technological Proficiency (1)	

End of Block: Default Question Block

Chapter 5

Machine learning algorithms enhance the accuracy of radiographic diagnosis of dental caries: a comparative study

This chapter presents a comparative study investigating the influence of cognitive aids such as machine learning (ML) algorithms and checklists on the accuracy of diagnosing dental caries on bitewing radiographs.

Title:

Machine Learning Algorithms Enhance the Accuracy of Radiographic Diagnosis of Dental Caries: A Comparative Study

Type of manuscript:

Original research

Full names of the authors:

1. Shwetha Hegde BDS, MDS, FHEA
   Academic Fellow,
   Dentomaxillofacial Radiology,
   Sydney Dental School,
   University of Sydney, Australia

2. Jinlong Gao, BDS, PhD, FHEA
   Associate Professor, Sydney Dental School,
   Institute of Dental Research,
   Westmead Centre for Oral Health,
   University of Sydney, Australia

3. Shanika Nanayakkara, MBBS, PhD, FHEA
   Senior lecturer, Sydney Dental School,
   Institute of Dental Research,
   Westmead Centre for Oral Health,
   University of Sydney, Australia
4. Stephen Cox, BDS, PhD
   Associate Professor, Head of Discipline,
   Discipline of Oral Surgery,
   Sydney Dental School,
   University of Sydney, Australia

5. Rena Logothetis, PhD
   MIEAust CPEng NER APEC Engineer IntPE(Aus)
   Senior Research Fellow,
   Applied Artificial Intelligence Institute,
   Deakin University, Melbourne, Australia
6. Rajesh Vasa, PhD
   Professor, Head of Translational Research and Development,
   Applied Artificial Intelligence Institute,
   Deakin University, Melbourne, Australia

Name, address, and email address of the corresponding author:

Dr Shwetha Hegde
Room 7.20, Level 7, Sydney Dental Hospital,
1, Chalmer Street, Surry Hills, NSW 2010
shwetha.hegde@sydney.edu.au

Conflict of interest statement:

The authors have no conflicts of interest or financial interest to declare.

Funding:

None

Acknowledgements:

We acknowledge www.denti.ai for providing access to their machine learning powered caries detection software.

Machine Learning Algorithms Enhance the Accuracy of Radiographic Diagnosis of Dental Caries: A Comparative Study

Abstract

Objectives:

This study evaluated the influence of cognitive aids, including machine learning (ML) algorithms and checklists, on the diagnostic accuracy and confidence of dental students in detecting dental caries on bitewing radiographs.

Methods:

Fifty-two third-year dental students were randomly assigned to control, ML, or checklist groups. The participants recorded their caries diagnoses (charting) on ten bitewing radiographs with a rating of confidence in the diagnosis. Diagnostic accuracy and reliability were measured and compared between the study groups for caries detection (present/absent). The inter-rater reliability of radiographic caries grading using the International Caries Detection and Assessment System II (ICDAS II) was assessed using weighted kappa. Participants also completed questionnaires on their perceptions of cognitive aids.

Results:

ML group showed the highest diagnostic accuracy and confidence levels. For caries detection, the ML group achieved the highest sensitivity (79%) and diagnostic odds ratio (DOR) (20.3), while the checklist group had the highest specificity (90.9%). The control group showed moderate sensitivity (67.9%) but outperformed the checklist group in this metric. Interrater agreement for caries detection was highest in the ML group (κ=0.702, 95% CI:0.692-0.713; $p<0.001$), followed by the checklist group (κ=0.604, 95%CI: 0.593-0.615; $p<0.001$) and the control group (κ= 0.590, 95% CI: 0.578-0.601; $p<0.001$). The weighted kappa for ICDAS II grading was highest in the ML group (κ=0.520, $p<0.001$), followed by the control group and the checklist group. Confidence levels differed significantly between groups ($p<0.001$), with the ML group reporting highest confidence.

Conclusion:

ML algorithms enhance diagnostic accuracy and confidence, possibly by reducing cognitive load. While standardising the diagnostic process, checklists were less effective, likely due to their lack of flexibility and clinical context. ML algorithms as cognitive aids are valuable tools for dental education and clinical practice. Further research is needed to better understand our findings and translate them into clinical workflows.

Clinical significance:

Machine learning (ML) algorithms improve the accuracy of radiographic diagnoses of dental caries and boost clinician confidence in diagnosis by reducing cognitive load. Their ability to support decision-making may be particularly valuable in complex cases, where they can help standardise radiographic interpretation, reduce inter-observer variability and improve overall diagnostic consistency. Integrating ML into clinical workflows could optimise diagnostic efficiency and enhance patient care outcomes.

Keywords: dental caries diagnosis, bitewing radiography, cognitive aids, machine learning, checklists

Introduction:

The diagnostic process relies on the dentist's ability to detect subtle differences between normal and abnormal findings in images and apply their knowledge to make accurate diagnoses.[1] Accurately detecting abnormalities on radiographs can lead to early detection of conditions and timely interventions, ultimately improving patient outcomes. However, the fast-paced and high-pressure nature of clinical environments increases the likelihood of errors and results in misdiagnosis and inappropriate treatment plans.[2] This may occur due to cognitive overload that impairs clinical decision-making.[3,4]

Cognitive load refers to the mental effort needed to complete a task or achieve a learning goal.[5] The cognitive load theory (CLT) describes how humans process new information by forming cognitive schemas- mental structures that help organise and interpret information- such as those used to systematically analyse a radiographic image.[6,7] CLT outlines three key components: i) memory systems, which include sensory, working and long-term memory; ii) processes involved in learning; and iii) various types of cognitive load that affect working memory.[8] According to the CLT, working memory is finite and can temporarily hold new information necessary for complex tasks, whereas long-term memory is more stable and organised, storing memory for extended periods and enabling more efficient information processing.[9] In situations of cognitive overload, the demands of working memory exceed capacity, and clinicians may miss critical information, leading to diagnostic errors.[10] For example, errors may occur when a clinician is presented with unfamiliar radiographic patterns or when the radiographic findings do not correlate with the clinical information.[11] These errors are relevant in dentistry, where accurate interpretation of dental radiographs is essential for diagnosing conditions like dental caries.

Dental caries remains a prevalent dental problem globally, and general dentists play a central role in the diagnosis and management of caries.[12,13] Diagnosing dental caries relies on clinical and radiographic examinations, and intraoral bitewing radiographs are a valuable diagnostic tool.[14–16] Studies investigating the diagnostic efficacy of dental imaging techniques in caries diagnosis have reported that intraoral bitewing radiographs are better than periapical, extraoral, and panoramic radiographs.[17] The radiographic diagnosis of caries is influenced by perceptual and cognitive factors, the clinician's experience, expertise, and results in variability in the diagnosis

of location and extent of caries.[18–21] The accuracy of caries diagnosis can also be influenced by the number and distribution of carious lesions visible on a radiograph, as previous research has demonstrated that the likelihood of detecting caries increases when more lesions and restorations are already present in the dentition.[22] False positive and false negative diagnoses and inter-observer and intra-observer variations further affect the diagnostic accuracy of bitewing radiographs.[23] The variations in radiographic caries diagnosis are also influenced by the location and depth of carious lesions, with higher accuracy observed for diagnosing dentinal caries on interproximal surfaces. [24] Initial caries on bitewing radiographs have low sensitivity and reasonable specificity, whereas deeper dentinal caries on interproximal surfaces have higher sensitivity and specificity. [25] These discrepancies in radiographic diagnosis can result in undertreatment or overtreatment of carious lesions. [26–28]

To address these challenges and improve the accuracy of caries diagnosis, various techniques have been developed, including templates, checklists, clinical decision support systems, and machine learning algorithms. [29,30,31] These techniques can serve as cognitive aids by simplifying complex tasks, assisting in decision-making by reducing mental effort and facilitating improved decision-making.[32–34] These cognitive aids have been tested in different clinical settings in medicine and have been shown to reduce errors, increase adherence to best practices and improve overall performance. Research has indicated that using guided checklists increases the accuracy of medical diagnoses and reduces errors in radiographic interpretation. However, their impact on diagnostic accuracy has not been measured quantitatively.[35,36] Checklists (CL) are underutilised in dentistry, and to the best of our knowledge, no studies have assessed their influence on the accuracy of radiographic caries diagnosis. In contrast, machine learning (ML) algorithms using deep learning techniques have been used to assist in the radiographic detection of dental caries.[30,37,38] Several studies have shown that algorithms increase diagnostic accuracy and reduce interpretation errors, with diagnostic accuracy between 80 and 90%. [30,38]

This study aimed to evaluate the influence of cognitive aids on the accuracy of diagnosing caries using bitewing radiographs. Specifically, the study aimed to compare the diagnostic accuracy of three groups of dental students: those using no aids, those using a diagnostic checklist and those assisted by an ML caries diagnosis algorithm. Additionally, the study aimed to evaluate the participant's self-reported confidence in their caries diagnosis and their opinions and perceptions about ML algorithms and checklists on diagnostic performance.

Materials and methods:

Study design and participant selection:

Dental students in their third academic year in the Doctor of Dental Medicine (DMD) program at Sydney Dental School, University of Sydney, were invited to participate in this study. The third-year dental students have a sound theoretical understanding of caries identification and the International Caries Detection and Assessment II (ICDAS II)[39] grading. Third year dental students had a strong foundational knowledge of caries diagnosis, having completed approximately 90

hours of didactic and pre-clinical training in cariology during the first two years of the program. From the second year onward, they began seeing patients in a supervised clinical setting. Their radiology training included approximately 15 hours of clinical experience where they actively took and interpreted intraoral and extraoral radiographs, completed structured radiographic assessments using ICDAS II and received feedback from clinical tutors. Additionally, during their clinical sessions in the student clinics, students routinely took and interpreted bitewing radiographs as part of the patient care with their findings reviewed by supervising clinicians. These combined experiences ensured that students had the necessary foundational knowledge and practical competence in caries diagnosis to participate in this study. However, no formal calibration for ICDAS II grading was conducted between the students and the clinical tutors specifically for this study. Additionally, while some participants may have had prior clinical experience (e.g., working as a dental assistant), which could introduce potential bias, this was mitigated in several ways. First, all participants received similar training in radiographic diagnosis of dental caries as a part of the DMD curriculum, ensuring comparable foundational knowledge. Furthermore, random allocation to the study groups helped control for individual differences in prior experience.

By recruiting dental students, who generally have similar levels of clinical experience, clinical experience was excluded as a confounding factor in diagnostic accuracy. This approach ensured that variations in diagnostic performance were not due to differences in clinical experience. This study received ethics approval from the University of Sydney's Human Research Ethics Committee (approval no. 2023/019). Based on the sample size calculation, each of the three groups- control, machine learning (ML) and checklist (CL), was required to have 16 participants. This calculation was made using a confidence interval of 95% and power of 80% with a 1:1:1 group ratio and expected variation in success rates of 60%, 70% and 80% (standard deviation of 10%).

Bitewing radiograph selection and quality assessment:

For this study, the ten highest quality adult bitewings were selected from the private collection of 50 de-identified bitewings belonging to the research team. Bitewing radiographs showing posterior teeth from the distal surfaces of the first premolars to the mesial surfaces of the second molars were considered for quality assessment. Two researchers (SH and JG) conducted a quality assessment (QA) of the images to ensure that images with adequate contrast, brightness, and appropriate techniques were included and bitewings with significant film faults were excluded. The following quality assessment criteria were applied: QA1-technically perfect, QA2-diagnostically acceptable, and QA3-non-diagnostic.[40] Only images with QA1 and QA2 ratings were included. Those radiographs with cone cuts or overlapping interproximal surfaces were also excluded. Bitewings that met the inclusion criteria were allocated a reference number.

Gold standard consensus:

A panel of five experienced clinicians served as the reference standard for diagnosing caries on each bitewing radiograph. The expert panel included four experts in dentomaxillofacial radiology with clinical experience ranging from 20 to 50 years, holding positions as faculty members at

Sydney Dental School and radiology consultant roles outside of the University. The fifth expert was a general dentist with 35 years of clinical experience. Specifically, each panel member independently reviewed every bitewing and assigned a caries diagnosis for each tooth surface, noting the presence or absence of caries and, where relevant, assigning a score based on the ICDAS II system. The ICDAS II radiographic criteria describe caries severity based on the depth of caries in enamel and dentine (Code 0- no caries; Code 1- caries involved half the thickness of enamel; Code 2- caries involving the entire thickness of enamel; Code 3-caries extending just into dentine (DEJ); Code 4-caries involved outer 1/3 of dentine; Code 5- caries involving inner 2/3rd of dentine) (Figure 1). Occlusal caries were also included as bitewings, which serve as an important adjunct, although occlusal caries are primarily diagnosed clinically. Bitewings are particularly useful for detecting dentine involvement or lesions not clearly visible on clinical examination alone.

Areas of disagreement were identified by comparing the expert's diagnosis, and a consensus meeting was held to resolve these discrepancies. Fleiss's multi-rater kappa coefficient was used to assess the overall level of agreement between experts. The final consensus annotations were used as gold standard annotations evaluating the accuracy of caries diagnosis by the participants in the study, ensuring that both the presence/absence and specific caries diagnosis were clearly documented.

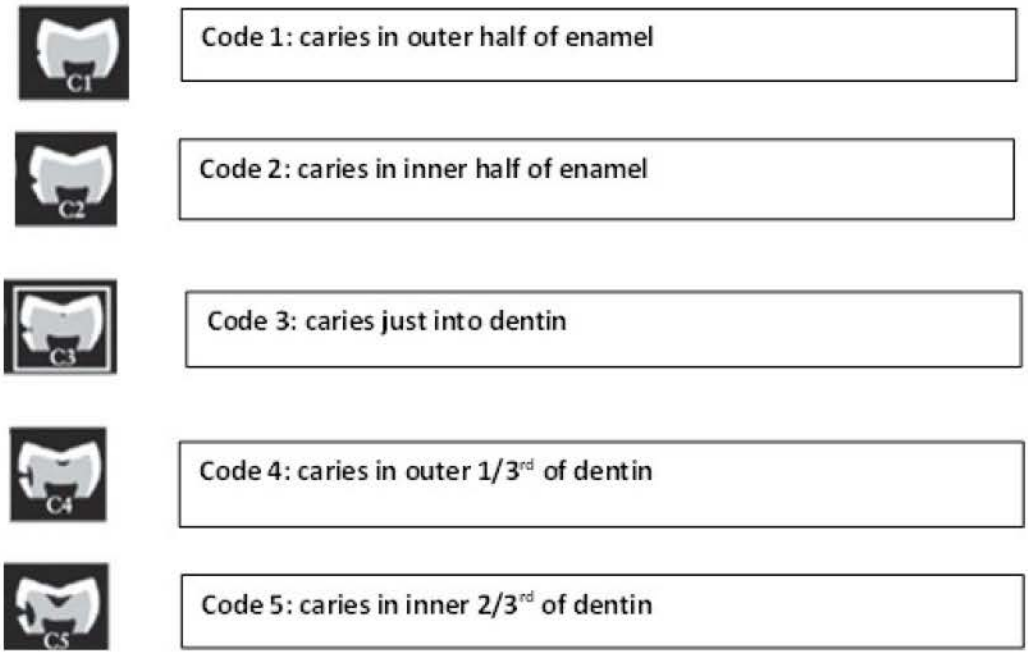


Figure 1: ICDAS II radiographic caries diagnosis criteria

Cognitive aids:

The research team developed a diagnostic checklist using international guidelines[40] to assist the participants in systematically analysing the bitewing for caries identification. The checklist is included in Supplementary File 1. From the web search results for an automated caries detection program, companies that marketed ML algorithms for caries detection were contacted, and www.denti.ai was chosen as they supported this study. Their ML algorithm is a deep learning based model trained on large datasets of annotated dental radiographs. While the specific proprietary details of the model architecture were not disclosed, denti. Ai has publicly reported the use of convoluted neural networks (CNNs), a widely used deep learning approach for image analysis in medical and dental applications. The training dataset for the denti.ai model consists of expert annotated dental radiographs, curated to include a wide range of carious and non-carious lesions. The model is continually refined through supervised learning, where radiographic images are labelled by experienced clinicians. The training process involves preprocessing steps, such as contrast enhancement and noise reduction, to ensure consistency in caries detection across different imaging conditions. The ML algorithm used in this study identified areas of suspected caries on bitewing radiographs but did not provide an ICDAS II grading to the participants in the ML group. Instead, it functioned as a binary (caries present/ absent) detection tool, assisting participants by highlighting regions of interest on the radiographs. The detection model was previously validated by internal testing and clinical studies conducted by denti.ai, demonstrating high sensitivity and specificity for caries detection. However, it is important to note that the model's performance may vary depending on image quality, variations in dental anatomy and operator interpretation.

Study procedure:

The study included ten bitewing radiographs, selected based on quality criteria and randomised (not presented in order of difficulty) to eliminate the cognitive impairment effect due to image order. The bitewing images were displayed on iPad devices using the Qualtrics platform, and all participants completed their analyses in the same environment with controlled ambient lighting. Image enhancement tools were not used in this study, ensuring that all participants assessed the radiographs under standardised viewing conditions without additional enhancements.

Participants were randomly assigned to three groups by a participant randomiser created using a Python script (Supplementary File 3), ensuring that each participant had an equal chance of being assigned to any group. The Python code was designed to randomly allocate participants (using the 'allocate_participants' function) into three groups, ensuring a 1:1:1 ratio. Each participant was given a participant reference number, and all collected data was anonymised. Group 1 was the control group, and the bitewings for caries were analysed without the help of a cognitive aid. Group 2 interpreted the same set of bitewing radiographs with the assistance of a caries detection machine learning algorithm (Supplemental Figure 1). Group 3 received the same images for caries interpretation and was assisted by a checklist.

For each bitewing radiograph, participants noted whether the caries were present or absent. The participant further classified caries using the 5-point ICDAS II grading, which was recorded on a paper-based diagnostic odontogram, shown in Figure 2.

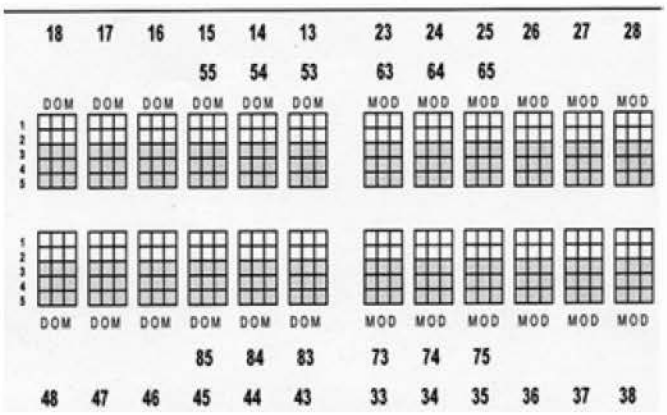


Figure 2: The diagnostic odontogram used to chart radiographic caries depth

Questionnaire:

At the end of the caries charting task, all participants completed an anonymised post-task questionnaire to record their feedback on their experience. The questions in the questionnaire varied depending on the group to which the participants were assigned. The questionnaire data was collected using Qualtrics (University of Sydney) (www.qualtrics.com), an online survey platform suitable for collecting and storing confidential and highly confidential data. All other data was stored in the University's OneDrive password-protected account.

The questionnaire was used to identify (i) the participant's views on how helpful the cognitive aids were in assisting with caries diagnosis and (ii) the participants' attitude towards cognitive aids—whether they would consider using them in future clinical practice.

The questionnaires are included in Supplementary File 2. The steps in the study are shown in Figure 3.

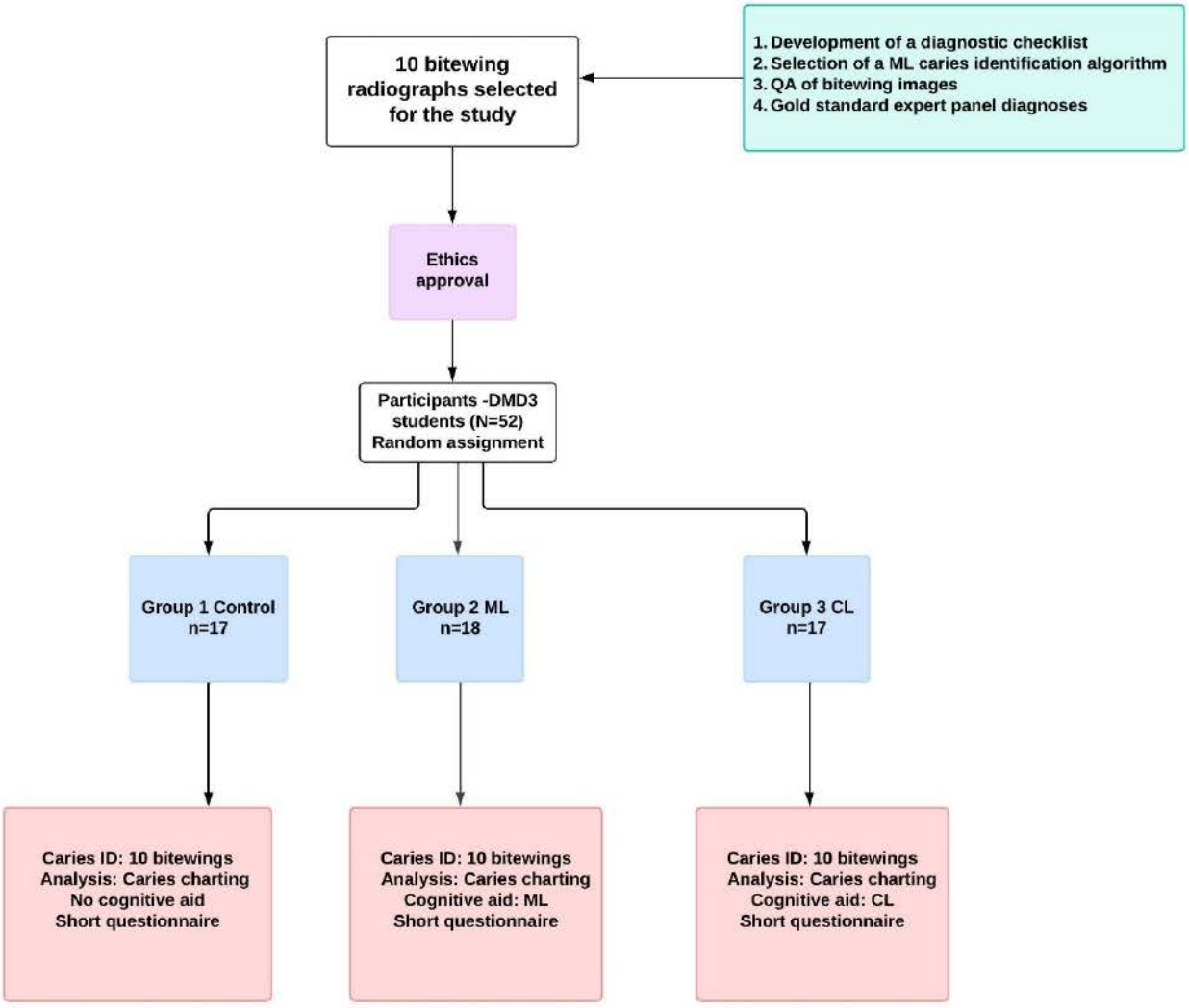


Figure 3: Flowchart illustrating the study procedures. (ML-machine learning; QA- quality assessment; DMD-Doctor of Dental Medicine; CL- checklist)

Data analysis:

Statistical analysis was performed using IBM SPSS statistics (IBM Corp. Released 2023. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp).

Descriptive statistics were used to summarise participant demographic data, including frequency, percentages and measures of central tendency and variability. The interquartile range method was used to remove any outliers in descriptive data records before analysis.[41] The chi-square test of independence was used to compare categorical data between the groups with Fisher's exact test when cell counts were low. For continuous variables such as age and the time taken for

radiographic diagnosis, which were not normally distributed, the non-parametric independent-samples Kruskal-Wallis test was used to compare those between the study groups.

Interrater agreement among the five experts who established the gold standard for caries detection and ICDAS II grading was assessed using Fleiss's kappa coefficient. The diagnostic performance for caries detection (caries present/caries absent) was evaluated using sensitivity (or recall; measures the ability to identify caries correctly), specificity (measures the ability to correctly identify the absence of caries), and diagnostic odds ratio (DOR; summarises the overall diagnostic performance) (Supplementary table S1). The interrater agreement between participants was compared for caries detection (caries present/absent) using Fleiss's kappa. The ordinal ICDAS II scores (caries charting) were analysed using weighted kappa to assess agreement between the participant's ICDAS II and the gold standard. Agreement between the three groups (control, CL, ML) was also compared using weighted kappa. The weighted kappa method accounts for ordinal data by assigning smaller penalties for minor disagreements (e.g., ICDAS II scores 1 versus 2) and larger penalties for major disagreements (e.g., ICDAS II scores 2 versus 5), ensuring that more severe misclassifications contribute more heavily to the overall measure of disagreement. The confidence ratings (measured on a 5-point Likert scale) were compared between groups using the Kruskal-Wallis test. Post hoc pairwise comparisons were performed using Dunn's test with a Bonferroni correction to identify significant differences.

The post-task questionnaire data were analysed using descriptive statistics.

Results:

Participant characteristics:

Fifty-two students enrolled in academic year 3 of the DMD program at Sydney Dental School participated in this study. The data was collected over 12 weeks from 27/07/2023 to 27/10/2023. Among the participants, there were 27 females (53.8%) and 24 males (46.2%). The participant's ages ranged from 23 to 36 years, with a mean age of 24.7 (±4.3) years. The mean age was comparable across the three study groups ($p>0.05$), with 23.6 (±6.4), 25.5 (±3.2) and 24.8 (±2.0), respectively, for the control, ML and CL groups.

On average, participants spent 23.1 (±6.2) minutes to analyse the ten bitewing images. The duration did not significantly vary across the study groups ($p>0.05$), with 21.6 (±5.6), 24.4 (±7.7) and 23.3 (±5.1) minutes, respectively, for the control, ML and CL groups.

Development of gold standard:

In the 10 bitewing images used in this study, 150 interproximal surfaces and 73 occlusal surfaces were visible. Of these surfaces, 56 had caries, and 167 surfaces were caries free. The images and details about the number and severity of caries (based on ICDAS II) are available in Supplementary File 4. The interrater agreement among the five experts who developed the gold standard was assessed using Fleiss's kappa showed substantial agreement ($\kappa=0.707$, 95% CI: 0.682-0.732; $p<0.001$) (Based on Landis and Koch's interpretation scale,[42] *<0.00 -poor agreement;*

0.00-0.20- slight agreement; 0.21-0.40- fair agreement; 0.41-0.60- moderate agreement; 0.61-0.80-substantial agreement; 0.8-1.00-almost perfect agreement).

Diagnostic performance of caries detection across study groups:

The diagnostic performance for dichotomous caries grading (caries present/absent) of the control, ML, and CL groups was evaluated using sensitivity, specificity and DOR (Table 1). The ML group achieved the highest sensitivity (79.1%), followed by the control group (67.9%) and the CL group (63.9%). In contrast, the CL group (90.9%) had the highest specificity, followed by the control group (87.6%) and the ML group (84.3%). The ML group had the highest DOR of 20.3, reflecting a strong discriminative power between positive and negative diagnoses. The CL group followed with a DOR of 17.7, and the control group had the lowest DOR at 15.0.

	Study Group					
Diagnostic test	**Control (n=17)**		**ML (n=18)**		**CL (n=17)**	
	Estimate	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**
Sensitivity (%)	67.9	64.9-70.7	79.0	76.5-81.4	63.9	60.9-66.9
Specificity (%)	87.7	85.9-89.4	84.3	82.4-86.2	90.9	89.3-92.4
Diagnostic odds ratio	15.0	12.2-18.6	20.3	16.5-24.9	17.7	14.1-22.3

Table 1: Diagnostic performance using dichotomous caries grading for ML, control and Checklist groups. (ML-machine learning; CI- confidence interval; CL- checklist)

Overall, the ML group showed superior performance on caries detection across several diagnostic metrics, including sensitivity and diagnostic odds ratio. The CL group had the highest specificity among the three groups but had lower scores for other diagnostic metrics compared to the ML group. When the performance of CL and the control groups were compared, the CL group had higher specificity and DOR compared to the control group. However, the control group performed better in sensitivity.

Interrater agreement for caries diagnosis:

The participants' interrater agreement was assessed at two thresholds. Interrater reliability for the caries detection (caries present/absent) within each group was assessed using Fleiss's kappa. ICDAS II scores were compared to the gold standard ICDAS II grading, and interrater reliability was evaluated using the weighted Cohen's Kappa (κ) coefficient.

For the caries detection task, the ML group demonstrated the highest agreement among participants (κ=0.702, 95% CI:0.692-0.713; p<0.001), indicating substantial agreement. The checklist group followed with a moderate agreement (κ=0.604, 95%CI: 0.593-0.615; p<0.001) and

the control group showed the lowest agreement (κ= 0.590, 95% CI: 0.578-0.601; p<0.001), also indicating moderate agreement.

For the ICDAS II scoring task, the ML group demonstrated the highest agreement with the gold standard (κ=0.520, p<0.001), followed by the control group (κ=0.476, p<0.001), and the checklist group showed the lowest agreement (κ=0.456, p<0.001). (*<0.00 indicates poor agreement; 0.00-0.20- slight agreement; 0.21-0.40- fair agreement; 0.41-0.60- moderate agreement; 0.61-0.80-substantial agreement; 0.8-1.00-almost perfect agreement*).

Confidence in caries diagnosis:

The confidence levels in caries diagnosis were recorded on a 5-point Likert scale for all ten bitewing images for the three groups, presented in Figure 4. A higher proportion of participants in the ML group reported greater confidence (quite confident/extremely confident) compared to the other two groups. Confidence levels varied across the different images for all groups, suggesting that some images were inherently more challenging or easier to interpret, regardless of the tool or method used.

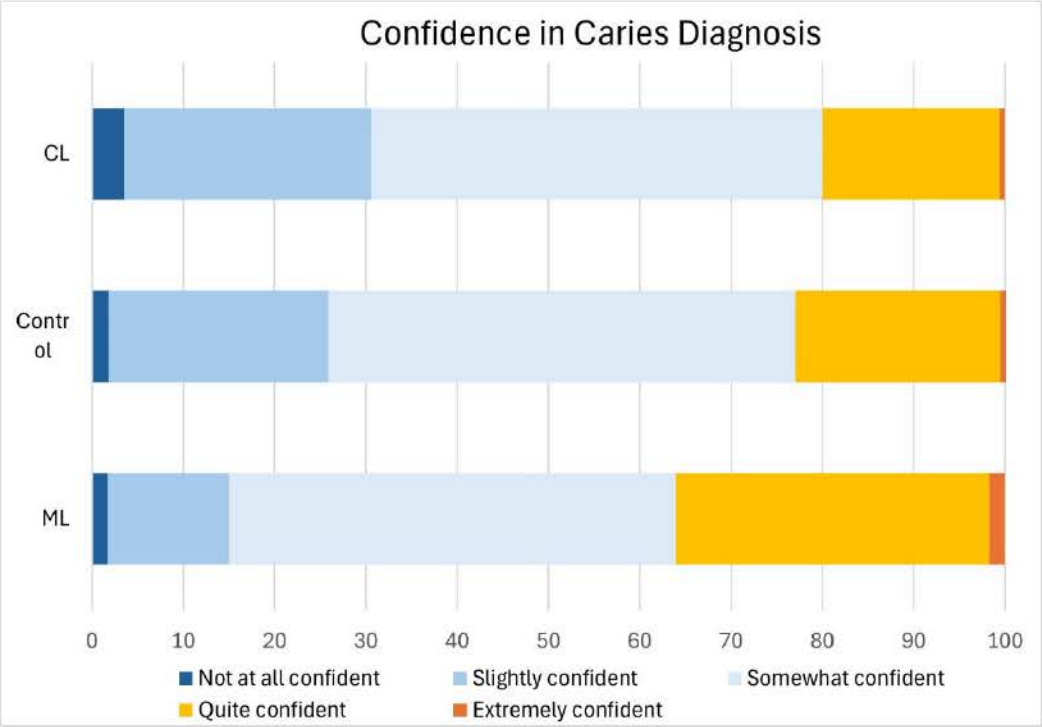


Figure 4: Confidence in caries diagnosis for the three groups

To compare confidence levels between the three groups, a Kruskal-Wallis test was performed, revealing a statistically significant difference between the three groups (H (2) =21.937, p<0.001). Post hoc Dunn's pairwise comparisons with Bonferroni correction showed the ML group had significantly higher confidence scores compared to the control (p<0.001) and checklist groups (p<0.001). However, there were no significant differences in confidence scores between the control and checklist groups (p=0.918).

Descriptive statistics were calculated to compare confidence scores across the three groups (control, checklist and ML). The ML group had the highest mean confidence score (M=3.23, SD 0.77), followed by the control group (M=2.96, SD 0.75) and the checklist group had the lowest mean confidence (M=2.86, SD 0.78). Median confidence scores were equal across all groups (Mdn=3, IQR 1), but the confidence interval of the ML group did not overlap with those of the checklist group, indicating meaningful differences in confidence levels.

A majority of the participants across the three groups reported being 'somewhat confident'. However, the higher mean score in the ML group suggests that a higher proportion of participants reported higher confidence levels (quite confident or extremely confident).

Post-task questionnaire regarding the use of cognitive aids in caries diagnosis:

Participants' views about their willingness to use the cognitive aids and their impact on diagnosis were assessed. Participants could choose more than one response to several questions.

In the control group (n=17), a majority (76.5%) of the participants showed interest in using checklists mainly to prevent missing lesions (58.8%) and to make a systematic assessment of the radiographs (41.2%). However, 23.5% in the control group were not inclined to use a checklist, citing reasons such as that it wastes time (17.6%), is inefficient (11.8%), or that there was no need for it (11.8%). When asked about the adoption of ML tools, 88.2% were open to considering ML tools in their practice. Of these, 35.5% indicated that ML could save them time. However, the majority of the participants (64.7%) did not specify their reasons. Only one participant (5.9%) was opposed to using ML tools in clinical practice.

In the ML group (n=18), participants varied in when they referred to ML prompts during their diagnostic process. One-third of the participants (33.3%) referred to ML prompts to assist them with the diagnosis before looking at the bitewing radiograph, 27.8% after analysis of the radiograph and 11.1% after viewing the radiograph to confirm the diagnosis. A small fraction of the participants checked the prompts but noted that it did not affect their diagnosis. Data was missing for 22.2% of the participants on this question. The primary reasons for using ML prompts were to validate the diagnosis (72.2%) and to obtain a second opinion (55.6%). Additionally, 44.4% indicated that ML prompts saved time, 16.7% expressed trust in ML, and 11.1% used them due to a lack of confidence in their initial caries assessment.

Regarding the agreement in diagnosis with the ML prompt, 50 % of the participants found that their caries diagnosis aligned with the ML prompt, while 27.8% indicated that they changed their diagnosis based on the ML prompt. Data was missing for 22.2% of the participants.

In the checklist group (n=17), participants also varied in when they referred to the checklist during their diagnostic workflow. 41.2 % of the participants used the checklists after viewing the radiograph, 29.4% before viewing the radiograph, and 11.8% after viewing to confirm their diagnosis. Data was missing for 17.6% of the participants. The reasons for using the checklist were primarily to feel more confident in their diagnosis (58.8%) and to enhance accuracy (11.8%). A significant majority of the participants (88.2%) did not provide reasons for using the checklist. When asked if the checklist was helpful, responses were evenly split, with 41.2% of the participants finding it helpful and 41.2% of the participants stating that it was not helpful, with data missing for 17.6% of participants.

Discussion:

This study compared the diagnostic accuracy and inter-rater reliability among three groups of dental students- control, machine learning (ML) and checklist (CL) groups at two thresholds: caries identification of bitewing radiographs (detection) and the more detailed ICDAS II grading of caries severity. The findings suggest that the ML assisted diagnosis improved overall diagnostic accuracy and confidence, highlighting the potential benefits of AI based decision support in radiographic analysis.

One of the key findings was that the ML enhanced diagnostic performance by reducing cognitive load. Radiographic interpretation can be a cognitively demanding task and requires clinicians to analyse complex visual data while maintaining diagnostic consistency. The ML tool likely functioned as a cognitive scaffold, allowing participants to focus their attention on high-risk areas, verify ambiguous cases and make more confident decisions. This is supported by the higher confidence levels reported by ML users, suggesting that ML tools can enhance decision making by providing real-time validation of clinical judgements. Future studies should employ objective measures to further investigate the specific cognitive processes involved, such as visual attention, decision latency or mental effort to better understand the mechanisms through which ML supports radiographic interpretation.

While the ML group demonstrated lower specificity compared to the checklist group, it achieved higher sensitivity, diagnostic odds ratio and confidence levels. The trade-off between sensitivity and specificity is an important consideration in diagnostic performance. Higher sensitivity means that ML is better at detecting caries, reducing the likelihood of missed lesions that could progress without intervention. However, this can come at the cost of increased false positives, where sound surfaces may sometimes be misclassified as carious. In populations with moderate to high prevalence of caries, such as the Australian adult population, where 1 in 3 adults have at least one untreated carious tooth surface,[43] specificity becomes particularly important to avoid

overdiagnosis and unnecessary treatment. Overdiagnosis can lead to unwarranted treatment, increased patient costs, and loss of sound tooth surface. Additionally, unnecessary treatment places an extra burden on already strained public healthcare resources, particularly in rural and remote areas, where access to dental care is limited. While ML-assisted diagnosis enhances sensitivity, further refinement of these algorithms is necessary to balance sensitivity and specificity. Importantly, reducing cognitive load and optimising specificity are not mutually exclusive goals. A clinician centered approach that integrates ML as a support tool rather than a replacement for human expertise ensures a balanced and effective diagnostic process. Future research should focus on ensuring that ML models effectively support accurate, efficient and resource conscious clinical decision making. In contrast, the checklist group exhibited higher specificity, meaning it was more conservative in diagnosing caries, potentially reducing unnecessary treatment, but this came at the cost of lower sensitivity, potentially leading to missed diagnosis.

Interestingly, the control group had higher sensitivity than the checklist group, indicating that participants relying solely on their clinical judgements were more likely to classify surfaces as carious rather than missing potential lesions. In contrast, the checklist group may have faced challenges due to the initial learning curve associated with checklists.[44] Given their relative inexperience, our participants may not have fully adapted to incorporating checklists into their diagnosis in the study. Additionally, checklists can sometimes lead to a mechanical approach to diagnosis, or the clinician may decide to skim or not entirely follow the checklists.[45] The unfamiliarity of the checklist may have contributed to the poorer performance in the checklist group in our study. As a result, checklists may not have been as effective as expected in reducing cognitive load and improving diagnostic accuracy. However, studies have shown that the usefulness of checklists may be improved by using digital checklists and relatively uncomplicated designs based on minimalist principles.[46]

Confidence in diagnosis is crucial in clinical settings where diagnostic uncertainty may increase cognitive load and affect clinical decision-making.[47] By automating the analysis and interpretation of dental radiographs, ML algorithms can provide insights and decision support by highlighting areas of dental caries.[48,49] This can help clinicians focus their attention effectively and reduce the likelihood of missed diagnoses. As a result, ML algorithms can free up cognitive resources that clinicians can allocate to other critical aspects of patient care. This suggests that ML algorithms can act as a valuable second opinion, validating the clinician's diagnosis and reassuring them. This increased confidence is particularly beneficial in clinical settings where diagnostic uncertainty can impact decision-making and patient outcomes. While studies in medicine discuss the impact of ML on diagnostic performance,[50,51] the influence of ML algorithms on clinician confidence in diagnosis has not been studied, particularly in dentistry. Despite using a structured aid, the checklist group did not experience the same confidence boost, possibly due to the additional cognitive steps required to use checklists effectively.

The questionnaire data provided valuable insights into the participants' experiences and attitudes towards using cognitive aids in dental caries diagnosis. The ML group's responses highlight the potential of ML algorithms to enhance diagnostic confidence and accuracy, especially as a second-opinion tool. Overall, the questionnaire responses indicate the need for user-friendly cognitive aids that can cater to novice and experienced clinicians.

This study was not without limitations. The bitewing images included in this study had minimal technical issues (film faults), such as interproximal overlap. This may not reflect the technical variability commonly encountered in clinical practice. The selection of bitewings in this study may have improved the diagnostic performance for all groups, including the ML group. Future studies should evaluate ML performance in realistic settings by incorporating images with film faults commonly encountered in general dentistry. This study considered the impact of cognitive aids on the diagnosis of caries, and the participants did not have to provide treatment to the patients based on the diagnosis. In that sense, this study could be considered 'low stakes', and the participants' diagnostic decisions did not impact patients. Additionally, the study participants were third-year dental students with limited clinical experience. Less experienced participants may be more open to ML prompts due to their need for greater support, whereas experienced clinicians are likely to rely on their own expertise and judgement. While the results are promising, the relatively small sample size and selective nature of the study participants consisting of third-year dental students from a single dental school may limit the generalisability of the study findings. The study utilised a specific ML algorithm for caries detection. The performance and user acceptance of different ML tools may vary. In addition, participants self-reported confidence levels, introducing a subjective element to the data. Individual biases can influence self-assessment and may not accurately reflect true diagnostic confidence or competence.

Future studies will include larger sample sizes with diverse participants, clinicians with a range of clinical experience and multicentre studies. Further research is also required to compare the impact of ML algorithms on the diagnostic accuracy of a range of dental conditions on various types of dental imaging, such as panoramic and CBCT images. It would also be worthwhile to conduct such studies in clinical settings where the treatments are provided to patients based on the input and support provided by cognitive aids. Further, it is essential to test the effectiveness of cognitive aids under conditions of increased cognitive load, such as complex cases and time pressure.

Conclusion:

Cognitive aids, particularly machine learning algorithms, can enhance diagnostic accuracy and clinician confidence while diagnosing dental caries from bitewing radiographs. Integrating them into dental education and clinical practice can shape the diagnostic capabilities of current and future practitioners, leading to consistent and reliable diagnoses and ultimately enhancing patient care. Investing in comprehensive training ensures dental students and clinicians can effectively utilise AI tools to improve diagnostic accuracy by reducing cognitive load.

References:

1. Ruskin KJ, Stiegler MP, Rosenbaum SH. *Quality and Safety in Anesthesia and Perioperative Care*. Oxford University Press; 2016. https://books.google.com.au/books?id=oTrKDAAAQBAJ

2. Hakanen JJ, Bakker AB, Demerouti E. How dentists cope with their job demands and stay engaged: the moderating role of job resources. *Eur J Oral Sci*. 2005;113(6):479-487. doi:10.1111/j.1600-0722.2005.00250.x

3. Hegde S, Gao J, Vasa R, Cox S. Factors affecting interpretation of dental radiographs. *Dentomaxillofacial Radiol*. 2022;52(2):20220279. doi:10.1259/dmfr.20220279

4. Kirsh D. A Few Thoughts on Cognitive Overload. *Intellectica Rev l'Association pour la Rech Cogn*. 2000;30(1):19-51. doi:10.3406/intel.2000.1592

5. Schaffer C, Goldart E, Ligsay A, Mazwi M, Gallant S, Ehrmann D. Take a Load Off: Understanding, Measuring, and Reducing Cognitive Load for Cardiologists in High-Stakes Care Environments. *Curr Treat Options Pediatr*. 2023;9(3):122-135. doi:10.1007/s40746-023-00272-3

6. Plass JL, Moreno R, Brünken R. *Cognitive Load Theory*. Cambridge University Press; 2010. https://books.google.com.au/books?id=mFJe8ZnAb3EC

7. Seel NM. Schema(s) BT - Encyclopedia of the Sciences of Learning. In: Seel NM, ed. Springer US; 2012:2933-2936. doi:10.1007/978-1-4419-1428-6_3

8. Young JQ, Van Merrienboer J, Durning S, Ten Cate O. Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. *Med Teach*. 2014;36(5):371-384. doi:10.3109/0142159X.2014.889290

9. Van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ*. 2010;44(1):85-93. doi:https://doi.org/10.1111/j.1365-2923.2009.03498.x

10. Alyahya MS, Hijazi HH, Alolayyan MN, Ajayneh FJ, Khader YS, Al-Sheyab NA. The Association Between Cognitive Medical Errors and Their Contributing Organizational and Individual Factors. *Risk Manag Healthc Policy*. 2021;14:415-430. doi:10.2147/RMHP.S293110

11. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. *Am J Roentgenol*. 2013;201(3):611-617. doi:10.2214/AJR.12.10375

12. Jain N, Dutt U, Radenkov I, Jain S. WHO's global oral health status report 2022: Actions, discussion and implementation. *Oral Dis*. 2024;30(2):73-79. doi:https://doi.org/10.1111/odi.14516

13. Oral-health-and-dental-care-in-Australia. Australian Institute of Health and Welfare. Published 2023. Accessed September 19, 2024. https://www.aihw.gov.au/reports/dental-oral-health/oral-health-and-dental-care-in-australia

14. Wenzel A, Pitts N, Verdonschot EH, Kalsbeek H. Developments in radiographic caries diagnosis. *J Dent*. 1993;21(3):131-140. doi:10.1016/0300-5712(93)90022-I

15. Dove SB. Radiographic diagnosis of dental caries. *J Dent Educ*. 2001;65(10):985-990.

16. Hintze H, Wenzel A, Danielsen B, Nyvad B. Reliability of visual examination, fibre-optic transillumination, and bite-wing radiography, and reproducibility of direct visual examination following tooth separation for the identification of cavitated carious lesions in contacting approximal surfaces. *Caries Res*. 1998;32(3):204-209. doi:10.1159/000016454

17. Kamburoglu K, Kolsuz E, Murat S, et al. Proximal caries detection accuracy using intraoral bitewing radiography, extraoral bitewing radiography and panoramic radiography. *Dentomaxillofacial Radiol*. 2012;41(6):450-459. doi:10.1259/dmfr/30526171

18. Mileman PA, Van Den Hout WB. Comparing the accuracy of Dutch dentists and dental students in the radiographic diagnosis of dentinal caries. *Dentomaxillofacial Radiol*. 2002;31(1):7-14. doi:10.1038/sj.dmfr.4600652

19. Moro BLP, Novaes TF, Pontes LRA, et al. The Influence of Cognitive Bias on Caries Lesion Detection in Preschool Children. *Caries Res*. 2018;52(5):420-428. doi:10.1159/000485807

20. Lazarchik DA, Firestone AR, Heaven TJ, Filler SJ, Lussi A. Radiographic Evaluation of Occlusal Caries: Effect of Training and Experience. *Caries Res*. 1995;29(5):355-358. doi:10.1159/000262092

21. Cave V, Hutchison C. Does time pressure impact on dentists' diagnostic performance? *Evid Based Dent*. 2019;20(3):81-82. doi:10.1038/s41432-019-0043-4

22. White SC, Kaffe I, Gornbein JA. Prediction of efficacy of bitewing radiographs for caries detection. *Oral surgery, oral Med oral Pathol*. 1990;69(4):506-513. doi:10.1016/0030-4220(90)90387-8

23. Wenzel A. Radiographic modalities for diagnosis of caries in a historical perspective: from film to machine-intelligence supported systems. *Dentomaxillofacial Radiol*. Published online 2021:20210010. doi:10.1259/dmfr.20210010

24. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review. *J Dent*. 2019;91:103226. doi:https://doi.org/10.1016/j.jdent.2019.103226

25. Schwendicke F, Tzschoppe M, Paris S. Radiographic caries detection: A systematic review and meta-analysis. *J Dent*. 2015;43(8):924-933. doi:10.1016/j.jdent.2015.02.009

26. Wenzel A, Haiter-Neto F, Gotfredsen E. Risk Factors for a False Positive Test Outcome in

Diagnosis of Caries in Approximal Surfaces: Impact of Radiographic Modality and Observer Characteristics. *Caries Res*. 2007;41(3):170-176. doi:10.1159/000099314

27. Baelum V, Heidmann J, Nyvad B. Dental caries paradigms in diagnosis and diagnostic research. *Eur J Oral Sci*. 2006;114(4):263-277. doi:https://doi.org/10.1111/j.1600-0722.2006.00383.x

28. Espelid I, Tveit AB, Fjelltveit A, Fjelltveitc A. Variations among Dentists in Radiographic Detection of Occlusal Caries. *Caries Res*. 1994;28(3):169-175. doi:10.1159/000261640

29. Dayo AF, Wolff MS, Syed AZ, Mupparapu M. Radiology of Dental Caries. *Dent Clin North Am*. 2021;65(3):427-445. doi:10.1016/j.cden.2021.02.002

30. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*. 2018;77:106-111. doi:https://doi.org/10.1016/j.jdent.2018.07.015

31. ten Bosch JJ, Angmar-Mansson B. Characterization and validation of diagnostic methods. *Monogr Oral Sci*. 2000;17:174-189.

32. Greig PR, Zolger D, Onwochei DN, Thurley N, Higham H, Desai N. Cognitive aids in the management of clinical emergencies: a systematic review. *Anaesthesia*. 2023;78(3):343-355. doi:https://doi.org/10.1111/anae.15939

33. Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit*. *Crit Care Med*. 2004;32(10). https://journals.lww.com/ccmjournal/Fulltext/2004/10000/Eliminating_catheter_related_bloodstream.6.aspx

34. Krombach JW, Edwards WA, Marks JD, Radke OC. Checklists and Other Cognitive Aids For Emergency And Routine Anesthesia Care-A Survey on the Perception of Anesthesia Providers From a Large Academic US Institution. *Anesthesiol pain Med*. 2015;5(4):e26300. doi:10.5812/aamp.26300v2

35. Bruno MA, Walker EA, Abujudeh HH. Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction. *Radiographics*. 2015;35(6):1668-1676. doi:10.1148/rg.2015150023

36. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. *Acad Med*. 2011;86(3):307-313. doi:10.1097/ACM.0b013e31820824cd

37. Ali R Ben, Ejbali R, Zaied M. Detection and Classification of Dental Caries in X-ray Images Using Deep Neural Networks. *ICSEA 2016 Elev Int Conf Softw Eng Adv*. 2016;(c):223-227.

38. Hung M, Voss MW, Rosales MN, et al. Application of machine learning for diagnostic prediction of root caries. *Gerodontology*. 2019;36(4):395-404. doi:https://doi.org/10.1111/ger.12432

39. Evans RW, Dennison PJ. The Caries Management System: An evidence-based preventive strategy for dental practitioners. Application for children and adolescents. *Aust Dent J*.

2009;54(4):381-389. doi:10.1111/j.1834-7819.2009.01165.x

40. Whaites E, Drage N. *Essentials of Dental Radiography and Radiology*. Elsevier Health Sciences; 2013.

41. Walfish S. A review of statistical outlier methods. *Pharm Technol*. 2006;30. https://api.semanticscholar.org/CorpusID:114377662

42. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977;33(1):159-174. doi:10.2307/2529310

43. Welfare AI of H and. Oral-health-and-dental-care-in-Australia. Published 2024. Accessed February 10, 2025. https://www.aihw.gov.au/reports/dental-oral-health/oral-health-and-dental-care-in-australia

44. Treadwell JR, Lucas S, Tsou AY. Surgical checklists: a systematic review of impacts and implementation. *BMJ Qual & Saf*. 2014;23(4):299 LP - 318. doi:10.1136/bmjqs-2012-001797

45. Al-Khafaji J, Townshend RF, Townsend W, Chopra V, Gupta A. Checklists to reduce diagnostic error: a systematic review of the literature using a human factors framework. *BMJ Open*. 2022;12(4):e058219. doi:10.1136/bmjopen-2021-058219

46. Mastrianni A, Sarcevic A, Hu A, et al. Transitioning Cognitive Aids into Decision Support Platforms: Requirements and Design Guidelines. *ACM Trans Comput Interact*. 2023;30(3). doi:10.1145/3582431

47. Alam R, Cheraghi-Sohi S, Panagioti M, Esmail A, Campbell S, Panagopoulou E. Managing diagnostic uncertainty in primary care: a systematic critical review. *BMC Fam Pract*. 2017;18(1):79. doi:10.1186/s12875-017-0650-0

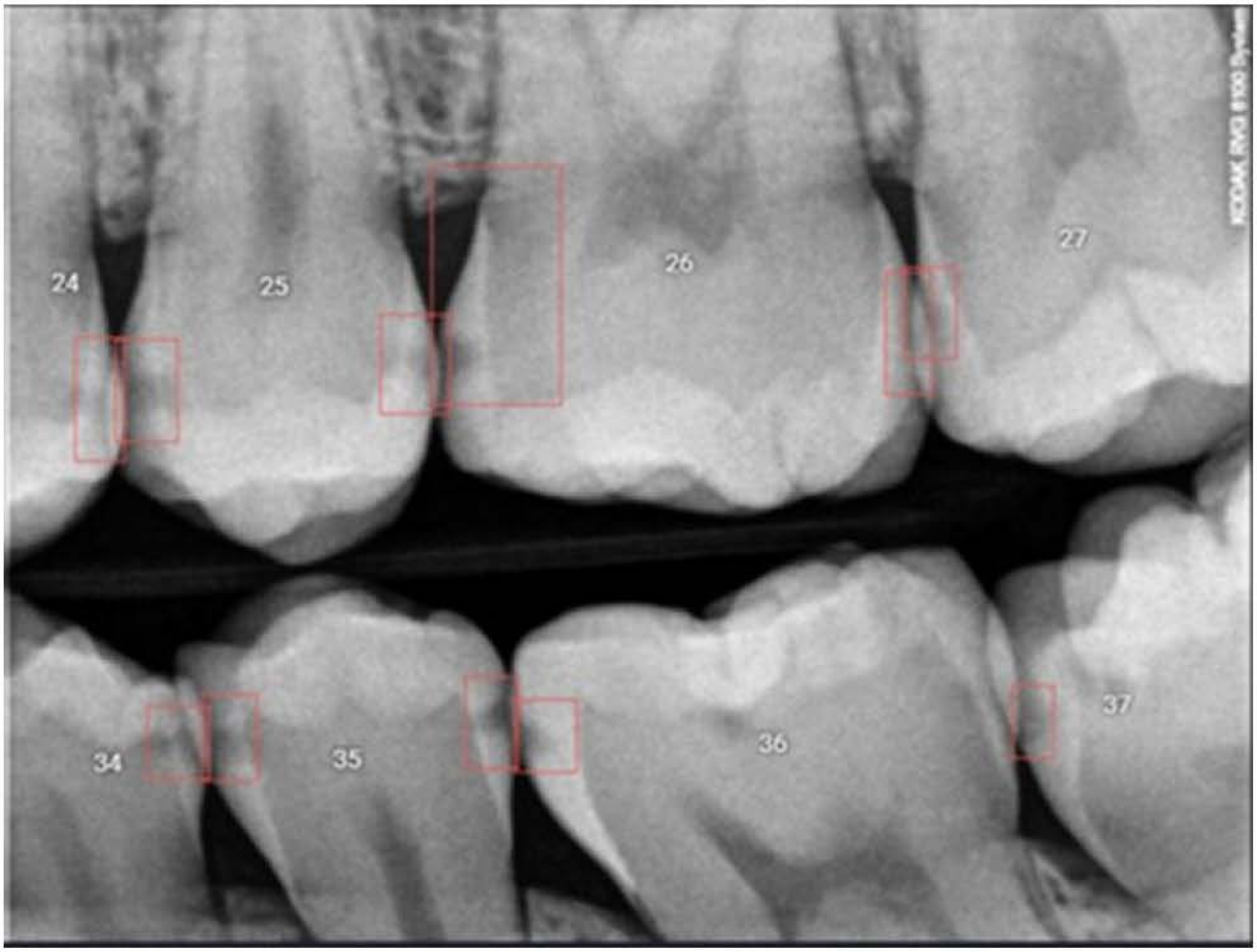
48. Albano D, Galiano V, Basile M, et al. Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review. *BMC Oral Health*. 2024;24(1):274. doi:10.1186/s12903-024-04046-7

49. Ammar N, Kühnisch J. Diagnostic performance of artificial intelligence-aided caries detection on bitewing radiographs: a systematic review and meta-analysis. *Jpn Dent Sci Rev*. 2024;60(November 2023):128-136. doi:10.1016/j.jdsr.2024.02.001

50. Grote T, Berens P. How competitors become collaborators—Bridging the gap(s) between machine learning algorithms and clinicians. *Bioethics*. 2022;36(2):134-142. doi:https://doi.org/10.1111/bioe.12957

51. Tschandl P, Codella N, Halpern A, et al. Human–computer collaboration for skin cancer recognition. *Nat Med*. 2020;26. doi:10.1038/s41591-020-0942-0

Supplementary Files:



Supplemental Figure 1: Bitewing radiograph with 'bounding box' caries identification provided by the ML algorithm (denti.ai).

Supplementary File 1:

Study Title: Influence of Cognitive Aids on Accuracy of Radiographic Diagnosis Of Dental Caries: A Comparative Study

Participant ID: _____

Date: _____

Please follow the steps outlined in the table to assist with your analysis of the bitewing.

1	Check the orientation of the radiograph	Right/left/premolar/ molar
2	Check image quality	QA1/QA2/QA3
3	Identify teeth seen on the bitewing.	
4	Are any radiolucencies indicative of caries?	Yes/No
5	If yes, assess from mesial to distal as if viewing the patient.	
6	Assess mesial-occlusal-distal surfaces of each tooth	
7	Determine the extension of caries using ICDAS coding	
8	Use the diagnostic odontogram to identify carious lesions.	

QA1- excellent quality image

QA2- diagnostic acceptable, has some film faults

QA3- non-diagnostic image

Supplementary File 2:

Post task questionnaire:

Post-task questionnaire/ Dental students

Survey Flow

Block: Default Question Block (6 Questions)

Branch: New Branch
If
If Which arm of the study did you participate in? Group 2/ML Is Selected

Block: ML (9 Questions)

EndSurvey:

Branch: New Branch
If
If Which arm of the study did you participate in? Group 1/ NA Is Selected

Block: NA (6 Questions)

EndSurvey:

Branch: New Branch
If
If Which arm of the study did you participate in? Group 3/CL Is Selected

Block: CL (6 Questions)

EndSurvey:

Page Break

Start of Block: Default Question Block

Q1 Please enter your participant ID

_____

Q21 Please state your age

_____

Q22 Gender

- ○ Male (4)
- ○ Female (5)
- ○ Non-binary / third gender (6)
- ○ Prefer not to say (7)

Q32 Please state your year of study

- ○ DMD 2 (1)
- ○ DMD 3 (2)
- ○ DMD 4 (3)

Q23 What were your undergraduate qualifications?

- ○ Oral health/dentistry (3)
- ○ Non oral health (4)

Q3 Which arm of the study did you participate in?

- ○ Group 1/ NA (1)
- ○ Group 2/ML (2)
- ○ Group 3/CL (3)

End of Block: Default Question Block

Start of Block: ML

Q8 Did you refer to the prompts generated by the machine learning (ML) algorithm to reach your caries diagnosis ?

- ○ Yes (1)
- ○ No (3)

Display This Question:

If Did you refer to the prompts generated by the machine learning (ML) algorithm to reach your carie... = Yes

Q24 When did you use the prompts provided by the ML algorithm in your diagnostic process?

- ○ Before looking at the radiograph, I checked the prompts to reach a diagnosis (1)
- ○ After looking at the radiograph, I checked the prompts to assist me with the diagnosis (2)
- ○ After looking at the radiograph, I checked the prompts to confirm my diagnosis (3)
- ○ After looking at the radiograph, I used the prompts to reach a diagnosis (4)
- ○ After recording my diagnosis, I checked the prompts, but it didn't affect my decisions (5)

Display This Question:

If When did you use the prompts provided by the ML algorithm in your diagnostic process? = Before looking at the radiograph, I checked the prompts to reach a diagnosis

Or When did you use the prompts provided by the ML algorithm in your diagnostic process? = After looking at the radiograph, I checked the prompts to assist me with the diagnosis

Or When did you use the prompts provided by the ML algorithm in your diagnostic process? = After looking at the radiograph, I checked the prompts to confirm my diagnosis

Or When did you use the prompts provided by the ML algorithm in your diagnostic process? = After looking at the radiograph, I used the prompts to reach a diagnosis

Or When did you use the prompts provided by the ML algorithm in your diagnostic process? = After recording my diagnosis, I checked the prompts, but it didn't affect my decisions

Q9 Why did you refer to the prompts given by ML algorithm? (multiple options may apply)

- [ ] It saves time (1)
- [ ] I was not confident to interpret the images (2)
- [ ] I trust the prompts made by the ML (3)
- [ ] To validate my diagnosis (4)
- [ ] As a second opinion (5)
- [ ] Other, please specify (6) _____

Display This Question:

If Did you refer to the prompts generated by the machine learning (ML) algorithm to reach your carie... = No

Q26 Why did you not refer to the prompts given by ML algorithm? (multiple options may apply)

- [ ] It wastes time (1)
- [ ] I was confident to interpret the images (2)
- [ ] I don't trust the prompts made by the ML (3)
- [ ] I did not like the user interface (4)
- [ ] I don't think the prompts by ML algorithm are accurate (5)
- [ ] Other, please specify (6) _____

Display This Question:

If Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = It saves time

Or Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = I was not confident to interpret the images

Or Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = I trust the prompts made by the ML

Or Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = To validate my diagnosis

Or Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = As a second opinion

Or Why did you refer to the prompts given by ML algorithm? (multiple options may apply) = Other, please specify

Q12 Was your caries diagnosis in agreement with the prompts by the ML algorithm?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Was your caries diagnosis in agreement with the prompts by the ML algorithm? = No

Q29 Did you change your diagnosis after seeing the ML prompt?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Did you change your diagnosis after seeing the ML prompt? = Yes

Q27 Why did you choose the ML prompt over your diagnosis?

_____

Display This Question:

If Did you change your diagnosis after seeing the ML prompt? = No

Q28 Why did you choose your diagnosis over the ML prompt?

_____

Display This Question:

If Was your caries diagnosis in agreement with the prompts by the ML algorithm? = Yes

Or Or Why did you choose the ML prompt over your diagnosis? Text Response Is Displayed

Or Or Why did you choose your diagnosis over the ML prompt? Text Response Is Displayed

Q13 Did the ML prompts affect your confidence in your diagnosis?

- ○ No (2)
- ○ Sometimes (3)

End of Block: ML

Start of Block: NA

Q15 Would you consider using a checklist to assist with caries identification?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Would you consider using a checklist to assist with caries identification? = Yes

Q16 If yes, please select your reason(s) from this list

- ☐ It saves time (1)
- ☐ It prevents lesions from being missed (2)
- ☐ Makes the image interpretation more systematic (3)
- ☐ Other (4) _____

Display This Question:

If Would you consider using a checklist to assist with caries identification? = No

Q33 If no, please select your reason(s) from the list

- [ ] It is inefficient (1)
- [ ] I don't need a checklist to help me (2)
- [ ] Wastes time (3)
- [ ] Other (4) _____

Q17 Would you consider using machine learning algorithm to assist you with caries identification?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Would you consider using machine learning algorithm to assist you with caries identification? = Yes

Q18 If yes, please select your reason(s) from the list

- [ ] It saves time (1)
- [ ] I am not confident to interpret the images (2)
- [ ] I trust the prompts made by the ML algorithm (3)
- [ ] To validate my diagnosis (4)
- [ ] As a second opinion (5)
- [ ] Other (6) _____

Display This Question:

If Would you consider using machine learning algorithm to assist you with caries identification? = No

Q34 If no, please select your reason(s) from the list

- ○ It wastes time (1)
- ○ I am confident to interpret the images on my own (2)
- ○ I don't trust the prompts given by the ML (3)
- ○ I don't the user interface (4)
- ○ I don't think the prompts by ML algorithm are accurate (5)
- ○ Other (6) _____

End of Block: NA

Start of Block: CL

Q2 Did you use the checklist to reach your caries diagnosis?

- ○ Yes (1)
- ○ No (3)

Display This Question:

If Did you use the checklist to reach your caries diagnosis? = Yes

Q4 When did you use the checklist?

- ○ Before looking at the radiograph (1)
- ○ After looking at the radiograph, I used the checklist to assist me in diagnosis (2)
- ○ After looking at the radiograph, I used the checklist to confirm my diagnosis (4)
- ○ After recording my diagnosis, I used the checklist but it didn't affect my diagnosis (5)

Display This Question:

If When did you use the checklist? = Before looking at the radiograph

Or When did you use the checklist? = After looking at the radiograph, I used the checklist to assist me in diagnosis

Or When did you use the checklist? = After looking at the radiograph, I used the checklist to confirm my diagnosis

Or When did you use the checklist? = After recording my diagnosis, I used the checklist but it didn't affect my diagnosis

Q6 Why did you use the checklist?

- ○ It saves time (1)
- ○ It makes me more confident my diagnosis (2)
- ○ It make the diagnosis more accurate (3)
- ○ Other (4) _____

Display This Question:

If Why did you use the checklist? = It saves time

Or Why did you use the checklist? = It makes me more confident my diagnosis

Or Why did you use the checklist? = It make the diagnosis more accurate

Or Why did you use the checklist? = Other

Q5 Did the checklist help you with the diagnosis?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If Did you use the checklist to reach your caries diagnosis? = No

Q31 Why did you not use the checklist?

- ○ It wastes time (1)
- ○ I am confident about my diagnosis (2)
- ○ It does not increase the accuracy of diagnosis (3)
- ○ Other (4) _____

Display This Question:

If Did the checklist help you with the diagnosis? = Yes

Or Why did you not use the checklist? = It wastes time

Or Why did you not use the checklist? = I am confident about my diagnosis

Or Why did you not use the checklist? = It does not increase the accuracy of diagnosis

Or Why did you not use the checklist? = Other

Q30 Did the checklist affect your confidence to diagnosing caries?

- ○ Yes (1)
- ○ No (2)

End of Block: CL

Supplementary File 3:

Python code

```
>>> import random
>>> participants = ["Participant1", "Participant2", "Participant3", "Participant4", "Participant5",
"Participant6", "Participant7", "Participant8", "Participant9", "Participant10", "Participant11" ]
>>> random.shuffle(participants)
... participants_per_group = len(participants) // 3
... group1 = participants[:participants_per_group]
... group2 = participants[participants_per_group:2 * participants_per_group]
... group3 = participants[2 * participants_per_group:]
... print("Group 1:", group1)
... print("Group 2:", group2)
... print("Group 3:", group3)
```

Supplementary File 4:

Gold standard:

The image set consisted of 10 bitewings with 150 interproximal surfaces and 73 occlusal visible. Of these, 56 surfaces had caries. Based on the gold standard diagnoses by experts the carious areas were classified as shown in the table below:

Caries classification criteria	Number of surfaces	Total surfaces
Code 0- no caries	162	162
Code 1- less than half of enamel	M-6; D-2; O-0	8
Code 2-more than half of enamel	M-11; D-6; O-0	17
Code 3- just into dentine	M-5; D-7; O-2	14
Code 4-outer one-third of dentine	M-2; D-5; O-4	11
Code 5- inner two-third of dentine	M-1; D-3; O-2	6
		56+162

Image wise caries on 10 bitewings

Image	Proximal surfaces visible	Occlusal surfaces visible	Caries	Caries free
1	18	8	6	20
2	16	8	4	20
3	12	6	1	17
4	12	7	1 (occlusal)	18
5	18	8	4+1 occlusal	21
6	20	8	8	20
7	12	7	10 +2 occlusal	7
8	16	8	6	18
9	12	6	6 +3 occlusal	9
10	14	7	3+1 occlusal	17
Total	150	73	56	167

Caries identification- gold standard consensus

Image number	Image type	Caries present
1	Right bitewing showing 18-13 and 44-47	18-Mcode 2; 17 D code 2; 16 D code 4; 15 M code 3; 46 Mcode 2; 47 M code 1
2	Left bitewing showing 23-27 and 34-37	26 D code 5; 27 M code 2; 35 D code 5; 37 M code 4
3	Right bitewing showing 17-14 and 44-47	16 M code 1
4	Left bitewing showing 24-27 and 34-37	27 occlusal code 3

5	Right bitewing showing 17-13 and 43-47	16 M code 5; 14 D code 5; 14 O code 5; 45 D code 2; 46 M code 1
6	Left bitewing showing 23-27 and 34-37	24 D code 2; 25 M code 1; 25 D code 2; 26 M code 2; 37 M code 1; 36 M code 2; 36 D code 1; 35 D code 1
7	Left bitewing showing 24-27 and 34-37	24 D code 4; 25 D code 3; 25 M code 3; 26 M code 2; 27 M code 2; 37 M code 2; 37 O code 4; 36 M code2; 36 O code 4; 35 M code 3; D code 3; 34 D code 4
8	Left bitewing showing 24-27 and 34-37	24 D code 2; 25 D code 4; 26 M code 2; 37 M code 3; 36 D code 3; 35 D code 3
9	Left bitewing showing 24-27 and 36-38	24 D code 3; 25 D code3; 25 M code 3; 26 M code2; 27 O code 4; 37 O code 3; 37 M code 4; 36 O code 5; 36 D code 4
10	Right bitewing showing 17-14 and 44-47	16 M code1; 15 D code 2; 45 D code 3; 47 O code 4

Image 1:



Image 2:



Image 3:



Image 4:



Image 5:

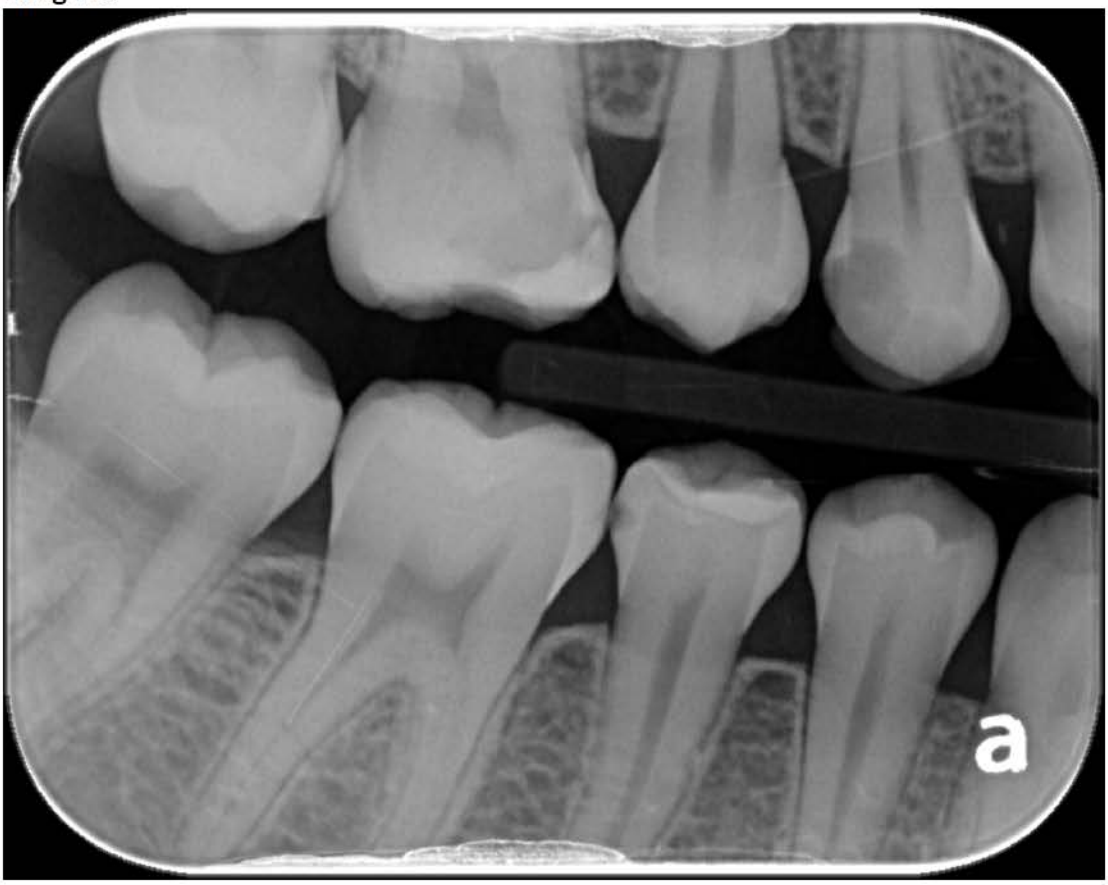


Image 6:

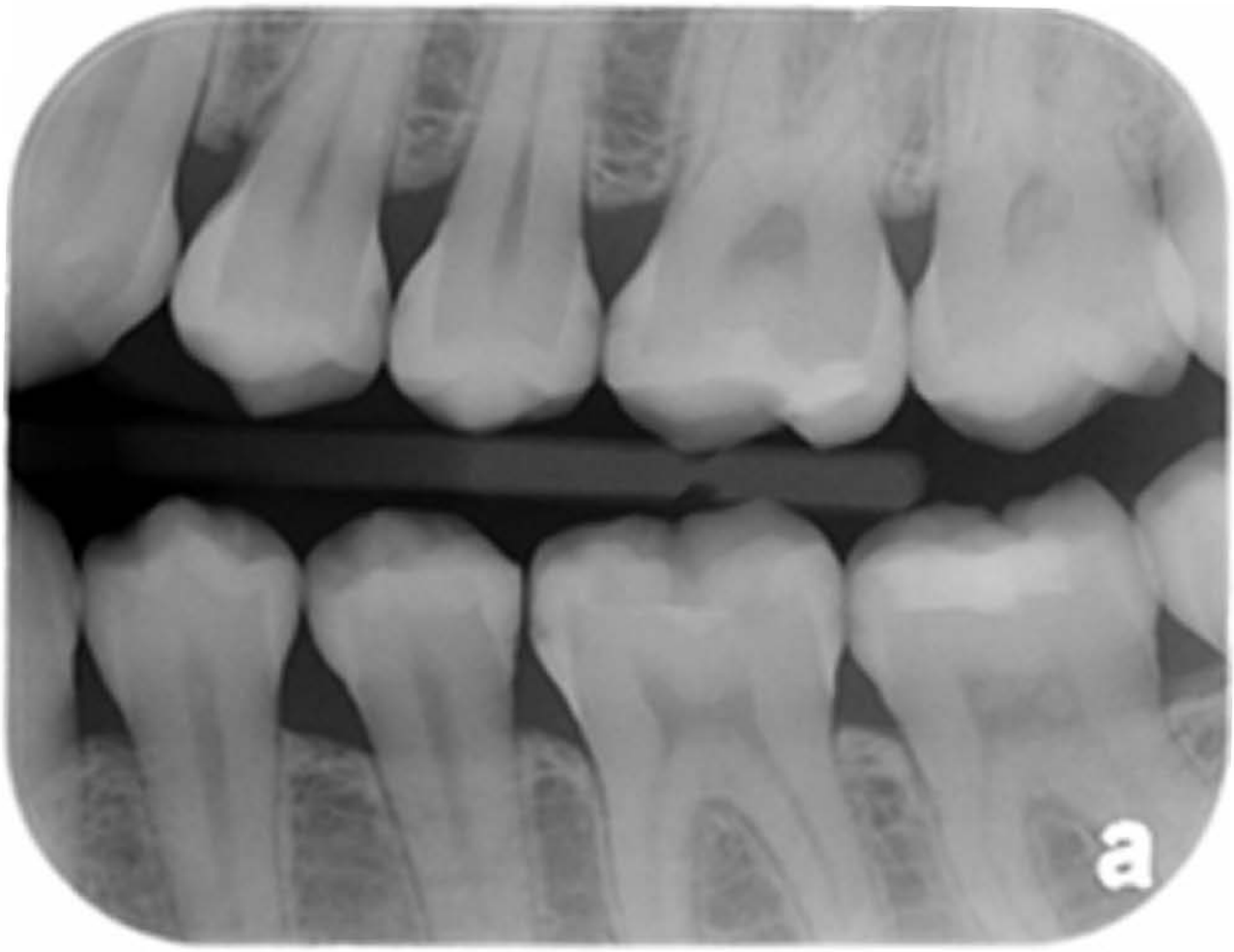


Image 7:



Image 8:



Image 9:

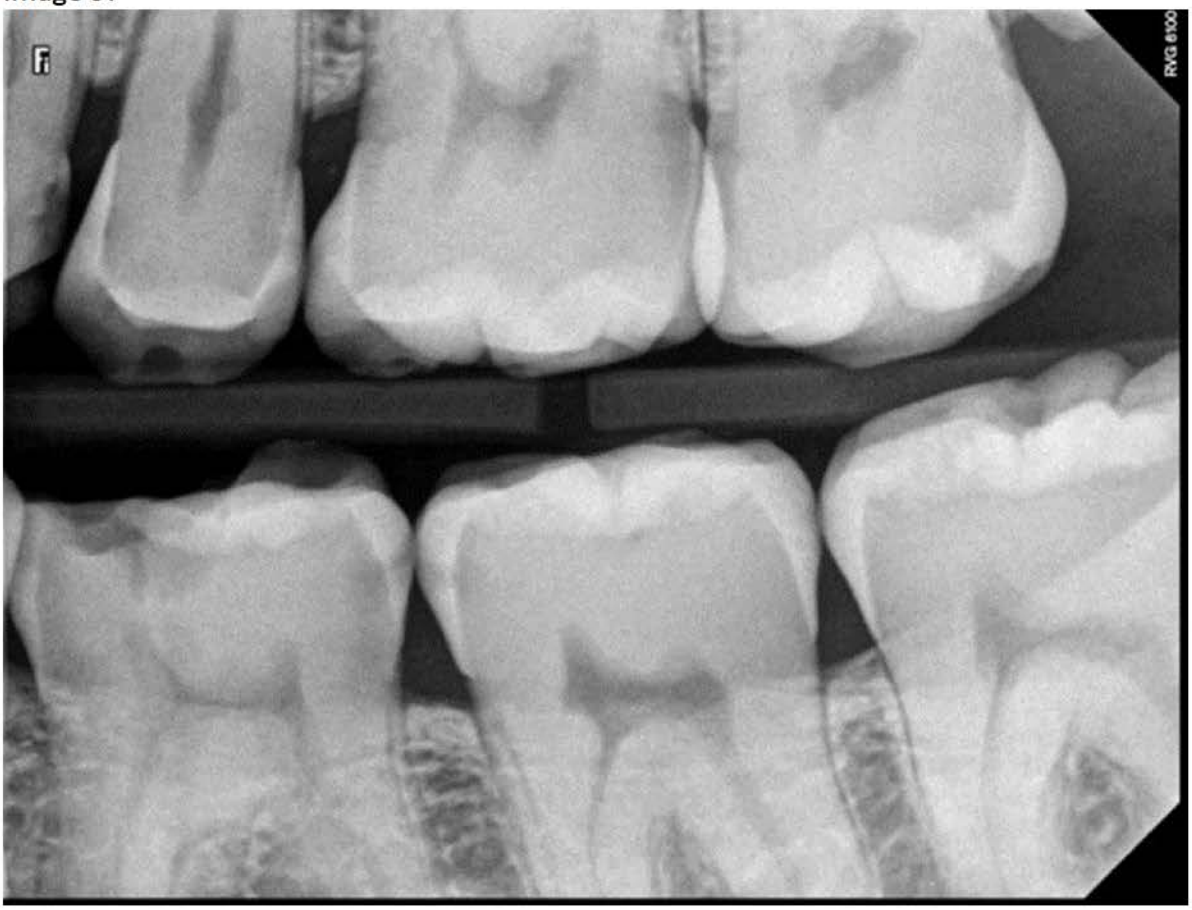


Image 10:



Supplementary Table

Diagnostic Metrics	Formula	Interpretation
Sensitivity (recall)	TP / (TP + FN)	The proportion of actual positives was correctly identified.
Specificity	TN / (TN + FP)	The proportion of actual negatives was correctly identified.
Diagnostic Odds Ratio	(TP / FN) / (FP / TN)	Odds of a positive test result being correct versus incorrect. Values >1 indicate a discriminatory ability; values>20 indicate strong discrimination.

Supplementary Table S1: Description of diagnostic metrics used in this study and their interpretation

Chapter 6

Attitudes and perceptions of Australian dentists and dental students towards applications of artificial intelligence in dentistry: A survey

Applications of artificial intelligence (AI) have sparked significant interest and research. Dentistry is no exception, with the potential of AI applications to improve diagnostic processes, treatment planning and patient care. This chapter presents a survey that explored the current perspectives within the Australian dental community regarding the role of AI in dentistry.

Check for updates

European Journal of Dental Education

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Attitudes and Perceptions of Australian Dentists and Dental Students Towards Applications of Artificial Intelligence in Dentistry: A Survey

Shwetha Hegde[1] | Shanika Nanayakkara[2] | Ashleigh Jordan[3] | Omar Jeha[3] | Usaamah Patel[3] | Vivian Luu[3] | Jinlong Gao[2]

[1]Dentomaxillofacial Radiology, Sydney Dental School, University of Sydney, Sydney, New South Wales, Australia | [2]Sydney Dental School, Institute of Dental Research, Westmead Centre for Oral Health, University of Sydney, Sydney, New South Wales, Australia | [3]Sydney Dental School, University of Sydney, Sydney, New South Wales, Australia

Correspondence: Shwetha Hegde (shwetha.hegde@sydney.edu.au)

Received: 15 May 2024 | **Revised:** 18 July 2024 | **Accepted:** 6 September 2024

Funding: The authors received no specific funding for this work.

Keywords: artificial intelligence | attitudes | dentistry | perception | survey

ABSTRACT

Introduction: As artificial intelligence (AI) rapidly evolves in dentistry, understanding dentists' and dental students' perspectives is key. This survey evaluated Australian dentists' and students' attitudes and perceptions of AI in dentistry.

Methods: An online questionnaire developed on Qualtrics was distributed among registered Australian dentists and students enrolled in accredited Australian dental or oral health programmes. Descriptive and bivariate analyses were used to examine the demographic variables and participant attitudes.

Results: 177 responses were received, and 155 complete responses were used in data analysis. 54.8% were aware of dental AI applications, but 70.3% could not name a specific AI software. A majority (91.6%) viewed AI as a supportive tool, with 69% believing that it would be beneficial in clinical tasks and 35.6% expecting it to perform similarly to an average specialist. 40% anticipated that dental AI would be routinely used in the next 5–10 years, with more dental students expecting this short-term integration. Concerns included job displacement, inflexibility in patient care, and mistrust of AI's accuracy. Attitudes towards AI were influenced by age, gender, clinical experience and technological proficiency.

Conclusions: The survey underscores the potential of AI to revolutionise dental care, enhancing clinical workflows and decision-making. However, challenges like trust in AI and ethical concerns remain. It is recommended that practising dentists receive hands-on training with AI tools and continuing dental education programmes. Integrating AI into dental curricula and fostering interdisciplinary teaching and research collaborations between computer science and dentistry is necessary to prepare graduates to use AI effectively and responsibly.

1 | Introduction

Artificial intelligence (AI) refers to machines that imitate human knowledge and behaviour [1]. The term artificial intelligence was first described by John McCarthy in 1955 [2]. In the last two decades, there has been a significant increase in research on applications of AI technologies in healthcare [3]. In 2023, AI in healthcare was worth 14.6 billion US dollars, and the market share is expected to be worth almost 102.7 billion US dollars by 2028 (https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-healthcare-market-54679303.html, https://www.statista.com/statistics/1334826/ai-in-healthcare-market-size-worldwide/). The

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *European Journal of Dental Education* published by John Wiley & Sons Ltd.

extensive use of electronic health records and digital imaging has facilitated the availability of huge data sets that has propelled the success of AI applications in healthcare [4, 5].

In clinical dentistry, AI technologies have been developed to aid in the radiologic diagnosis of different oral and maxillofacial pathologies and provide decision support and analysis of treatment outcomes in various dental disciplines [6–15]. Several commercially available AI technologies are helpful in routine dental care and reliable tools in providing a second opinion for image and patient data, improving clinical efficiency and thus saving valuable clinical time [16–22] (e.g., PearlAI [16] in caries detection). AI can process various data types, such as patient history, demographic information and treatment records, to help identify patient preferences and specific needs and boost motivation and accountability towards their health. AI-enabled smart assistants and chatbots can support dental consultations through teledentistry [23]. Despite these challenges, the future implications of AI are promising. AI has the potential to streamline clinical workflows, standardise diagnosis and treatment, increase the quality of clinical decisions, and improve patient safety by reducing errors [24]. AI can streamline administrative tasks in clinical practice, and this can free up time and resources for more patient-centred initiatives, making healthcare more accessible and affordable [25]. AI-enabled virtual and augmented reality (VR and AR) devices and gamification techniques can also assist with student training [26].

Although AI holds significant promise, its routine adoption in dental practice remains limited. Challenges that impede wider adoption include data accessibility, lack of replicability and robustness in dental AI research, and limited capabilities of current AI applications [5]. Additionally, there are concerns about data privacy and security, lack of generalisability of the algorithms, and inherent bias that may skew diagnostic or treatment recommendations [5, 27]. The opaque 'black-box' nature of the algorithm design has raised concerns about transparency and explainability in the decisions made by AI algorithms [28]. In addition, overreliance on technology may diminish the human element in dental care, impacting the quality and comprehensiveness of care delivered [23].

New avenues in AI research have been explored to address the concerns and translate research into effective clinical tools. This latest research includes understanding the interactions between clinicians and AI technologies. It also involves assessing the impact of AI technologies on the quality, efficiency and productivity of clinical practice and understanding how users adopt new technologies such as AI [29, 30]. User acceptance and experience can significantly influence attitudes and perceptions towards adopting new technology [30–32]. Based on the perceptions of potential users, evaluations have been conducted to identify usability issues, gaps in understanding, and possible barriers and concerns associated with AI applications in medicine and dentistry [33–38].

Evidence from recent literature suggests that the adoption of new technologies, particularly AI, can vary significantly among different healthcare settings, which are affected by the societal and cultural contexts and the clinical workflows [39–41]. Understanding the user experience within the Australian context enables the customisation of AI applications, promoting trust and acceptance among dentists and dental students. This understanding ensures that AI technologies align with local workflows, regulatory frameworks and patient expectations, maximising their effectiveness and utility [42].

Consequently, it is crucial to understand the perceptions and expectations of the Australian dental profession regarding AI technologies. This understanding will facilitate a better grasp of how and when these technologies will be adopted and implemented in routine dental care. This survey aimed to study the attitudes and perceptions of Australian dental practitioners and students about the applications of artificial intelligence in dental practice.

2 | Methods

A cross-sectional survey was designed and distributed through the online Qualtrics platform (www.qualtrics.com). The study received ethics approval (Human Research Ethics Committee approval number 2021/454) and was conducted in accordance with the Declaration of Helsinki [43]. The following inclusion and exclusion criteria were used to invite the participants:

Inclusion criteria:

1. Dentists, dental specialists and oral health therapists registered to practice in Australia
2. Dental and oral health students studying in any accredited Australian university

Exclusion criteria:

1. Dental practitioners other than dentists, dental specialists and oral health therapists
2. Dental practitioners not registered to practice in Australia

The dental students studying at Australian universities were contacted by an announcement in their University's online learning management systems. Dentists and dental specialists were contacted through the research team's known networks and passive snowball recruitment [44]. Social media platforms, including Facebook and LinkedIn, were also used to circulate the anonymised survey link. Participation was voluntary, and no incentives were offered. Informed consent was considered when participants completed and submitted the survey. The survey remained open from September 2021 until August 2022. Further data collection was stopped when regular monitoring and analysis indicated data saturation. This approach was adopted based on the grounded theory in qualitative research [45].

The survey consisted of an anonymised questionnaire containing multiple choice, Likert scale and open-ended questions. The research team developed the questionnaire based on the literature review and similar previously published surveys [33, 37]. The questionnaire was designed to present appropriate questions to dental practitioners and students using the branching option within Qualtrics. The questions covered

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the participants' awareness and perceptions about the current applications of AI algorithms in dentistry, their impact on dental workflow and the dental curriculum, potential benefits of AI-based solutions in dentistry, performance and accuracy of dental AI applications and concerns about adopting AI solutions in dentistry.

Demographic questions included participants' age, gender, professional qualifications (category), type of practice, years of clinical experience and location of practice (metropolitan/rural/private/public/academia). Student participants were asked to indicate the type (oral health or dental) and entry (undergraduate or postgraduate) of the programme they were enrolled in. The questionnaire was pilot tested to ensure that the questions were relevant, appropriate and accessible. The survey questionnaire is included in Data S1.

Data were transferred from the survey platform to IBM's SPSS statistics software (version 26, IBM, SPSS Inc., Chicago, IL) and analysed. Descriptive statistics summarised the demographic characteristics of the survey respondents. The association between the categorical variables was assessed using the chi-squared test of independence. Fisher's exact test was used when more than 20% of the cells had an expected frequency of less than 5. The Spearman rank correlation test was used to assess the correlation between variables. The level of significance for all tests was set at $p < 0.05$.

3 | Results

A total of 177 responses were received, and after removing incomplete records, 155 responses (87.6%) were used in the data analysis. Table 1 summarises the demographic characteristics of the participants. Of the 155 participants, 64 were dental students (43.2%), seven were oral health students (4.7%) and 77 were registered dentists (52%). The participants' ages ranged from 22 to 85 years (average of dental students 25 ± 3.7 years; dentists 37.8 ± 14.8 years) (Figure 1), with equal gender distribution. On average, the dentists practised 4.4 days per week, and most of the dentists practised in the private sector ($n = 42$, 27.1%) and in metropolitan areas ($n = 52$, 33.5%). The clinical experience of the dentists ranged from 1 to 43 years, with an average of 12.6 (± 11.8) years.

Among 146 valid responses, participants' self-assessment of their technological proficiency was 4.02 on a scale of 0–5 (Figure 2), with similar average scores for dental students and practitioners. Dentists used several digital technologies in their clinic, such as digital X-ray imaging (45.2%), digital patient records (43.9%) and OPG machines (38.7%).

3.1 | User Experience and Acceptance of Dental AI Applications in Clinical Practice

Among the survey participants, 54.8% ($n = 85$) were aware of AI applications in the field, and they mainly learned about dental AI applications through lectures, conferences, journals and social media (Table 2). Interestingly, a majority ($n = 109$, 70.3%) were unable to name a specific dental AI software, with a slightly higher proportion of student participants ($n = 59$, 54%) compared to dentists ($n = 47$, 43%) in this category. Among the 29.7% ($n = 46$) who could name a specific dental AI software, their knowledge primarily came from scholarly and clinical sources. The majority of participants identified dentomaxillofacial radiology (64.5%) and implantology (64.5%) as the most amenable disciplines for AI applications (Figure 3). This finding was augmented by the participant's preferences for AI use across various other areas, including diagnostics, general dentistry and emergency triage, suggesting a broad scope for practical integration of AI. AI was found to be most beneficial for tasks such as image processing and diagnosis and less beneficial for administrative tasks and treatment planning.

AI was recognised as a supportive tool for clinicians by 91.6% ($n = 142$) of the participants, and only 2% ($n = 3$) had concerns about potential negative effects. A majority of the participants ($n = 107$, 69%) indicated that AI would be beneficial to

TABLE 1 | Demographic data.

Participant's qualifications, *n* (%)	
Dental students	55 (35.5)
Oral health students	18 (11.6)
General dentists	68 (43.9)
Specialist dentists	11 (7.1)
Data missing	03 (0.9)
Participant's gender, *n* (%)	
Males	73 (47.1)
Females	73 (47.1)
Data missing	9 (5.8)
Participant's age (years)	
Mean (±SD)	32.3 ± 12.5
Minimum	18
Maximum	85
Years in clinical practice (years)	
Mean (±SD)	12.6 ± 11.8
Minimum	1
Maximum	43
Working days per week (mean ± SD)	4.4 ± 1.3
Type of practice (choice count)[a]	
Metropolitan	52 (33.5)
Rural	13 (8.4)
Public	15 (9.7)
Private	42 (27.1)
Academia	12 (7.7)

[a]The total does not add up to 100% as the participants chose more than one option.

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

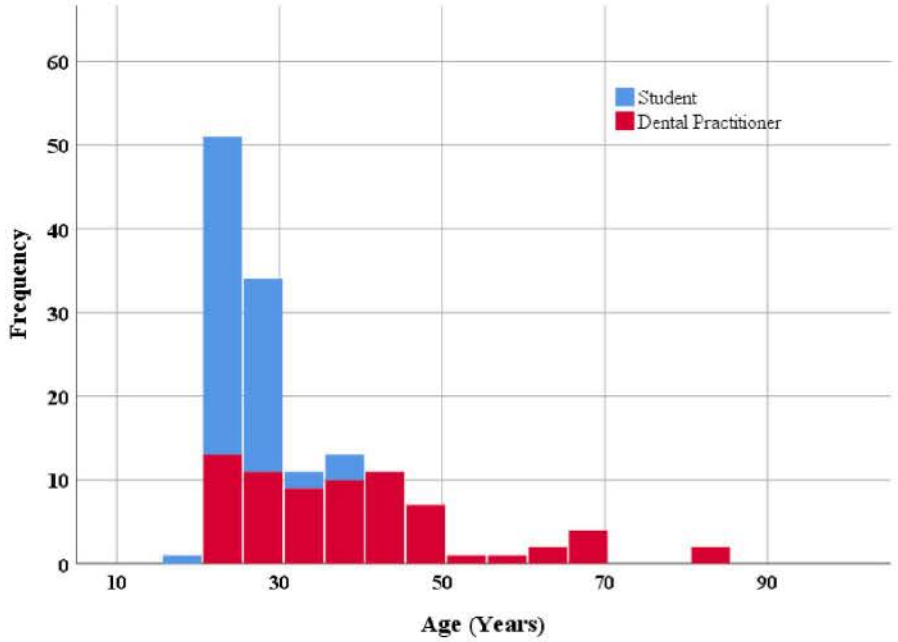


FIGURE 1 | Dental student and practitioner age range.

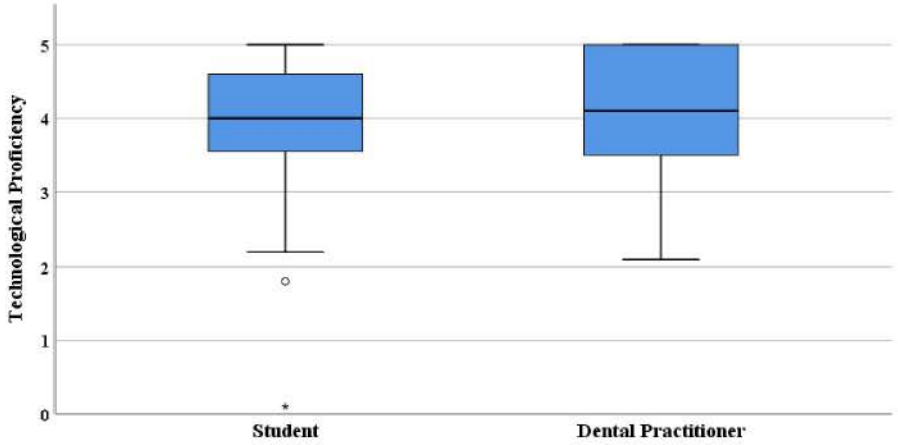


FIGURE 2 | Participant's self-assessment of technological proficiency.

TABLE 2 | Differences in the responses among dentists and dental students.

	Have you heard of AI in dentistry?			
Student or dental practitioner	**Yes, *n* (%)**	**No, *n* (%)**	**Total, *n* (%)**	***p***
Dental student	32 (45.1)	39 (54.9)	71 (100)	**0.023**
Dental practitioner	49 (63.6)	28 (36.4)	77 (100)	
Total	81 (54.7)	67 (45.3)	148 (100)	

Note: Bold value denote statistical significance at the $p < 0.05$.

clinical tasks in dentistry. However, 6% ($n = 9$) indicated that either AI would make no difference or have a negative impact. The participants revealed diverse opinions on the expected performance of AI compared to specialists: 35.6% ($n = 52$) expected AI performance to equal an average specialist, while 19.9% ($n = 29$) thought AI could outperform the best specialist. Additionally, 23.3% ($n = 34$) saw AI on par with the least effective specialist, 13% ($n = 19$) with the best and 8.2% ($n = 12$) predicted that AI would surpass the top specialists. There was no significant difference in expectations between dentist and student participants ($p > 0.05$).

Regarding the timeline for AI integration, 40% ($n = 60$) of all participants indicated that AI applications would be routinely

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

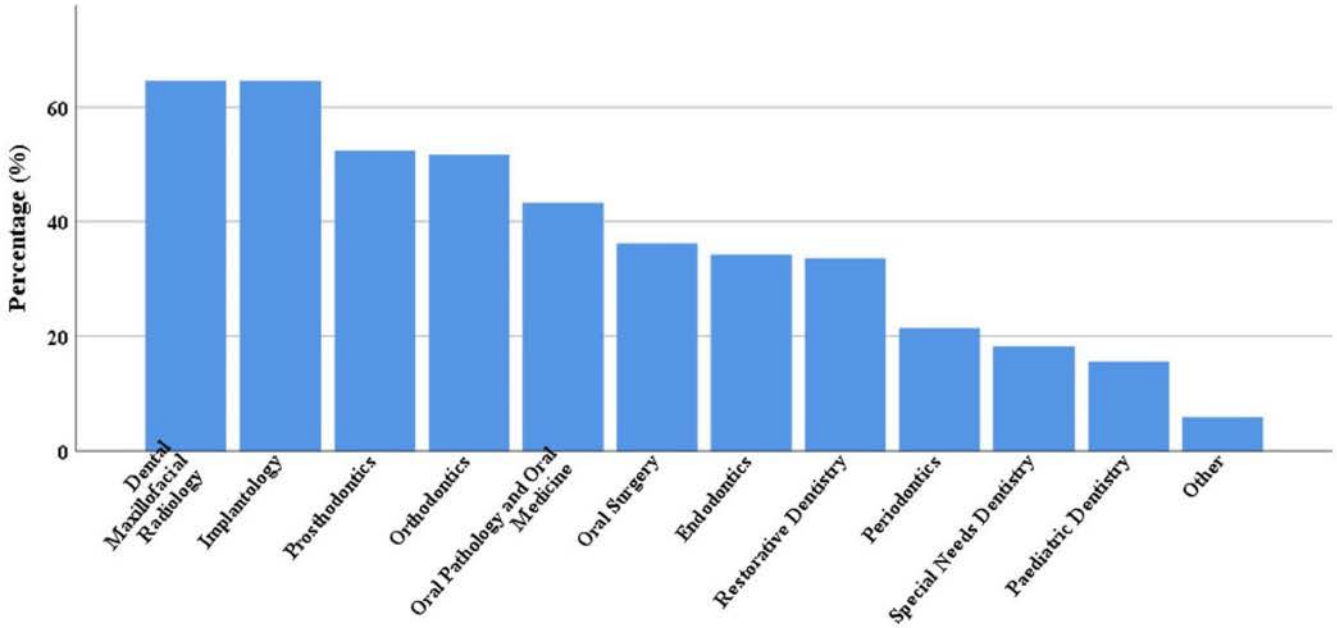


FIGURE 3 | Dental disciplines most likely to benefit from AI.

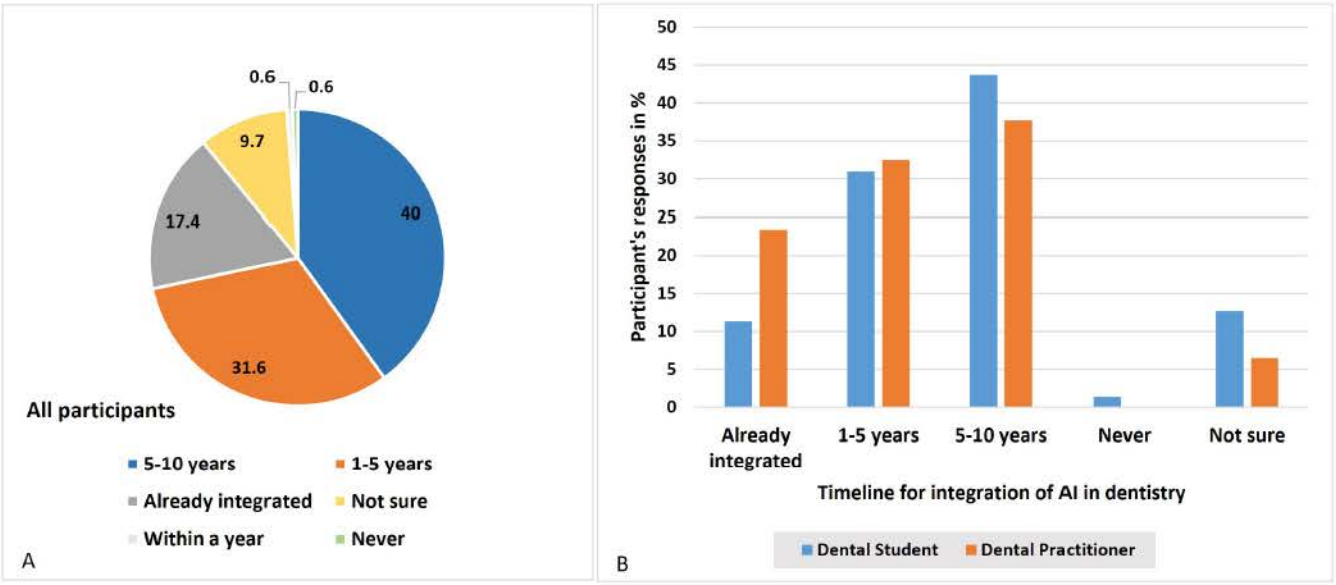


FIGURE 4 | Timeline for integration of AI in dentistry. (A) All participants. (B) Distribution of responses among dentists and dental students.

used in dentistry within the next 5–10 years (Figure 4), with similar opinions among dentists ($n = 29$, 48.3%) and dental students ($n = 31$, 51.6%). Interestingly, 23.4% ($n = 18$) of dentist participants, compared to only 11.3% ($n = 8$) of student participants, considered AI to be already integrated into dental practice. In comparison with dentists, dental students anticipated AI integration in clinical practice in the short term ($p = 0.022$). Over a third of dentist participants ($n = 53$, 34.2%) expressed excitement about integrating AI into their practice, whereas 17.4% ($n = 27$) believed AI would make little difference or had reservations. When faced with a discrepancy between AI's judgement and their own, a majority of participants, 59.6% ($n = 87$), indicated they would consult a colleague or an experienced clinician, whereas 8.9% ($n = 13$) indicated they would resort to other measures, such as searching scientific literature to resolve the differences (Figure 5). More students ($n = 46$, 68.7%) than dentists ($n = 35$, 48.6%) indicated they would refer to a colleague or a senior clinician. In contrast, more dentists ($n = 27$, 37.5%) indicated they would trust their own judgement compared to dental students ($n = 15$, 22.4%).

The survey also investigated the concerns about AI applications. The primary concerns identified were job losses to more efficient technology, lack of flexibility in patient care and insurance liability. Additionally, mistrust in the technology and concerns about its accuracy were highlighted. The survey found no statistically significant differences in the responses to concerns about AI applications among dentists and dental students.

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejn.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

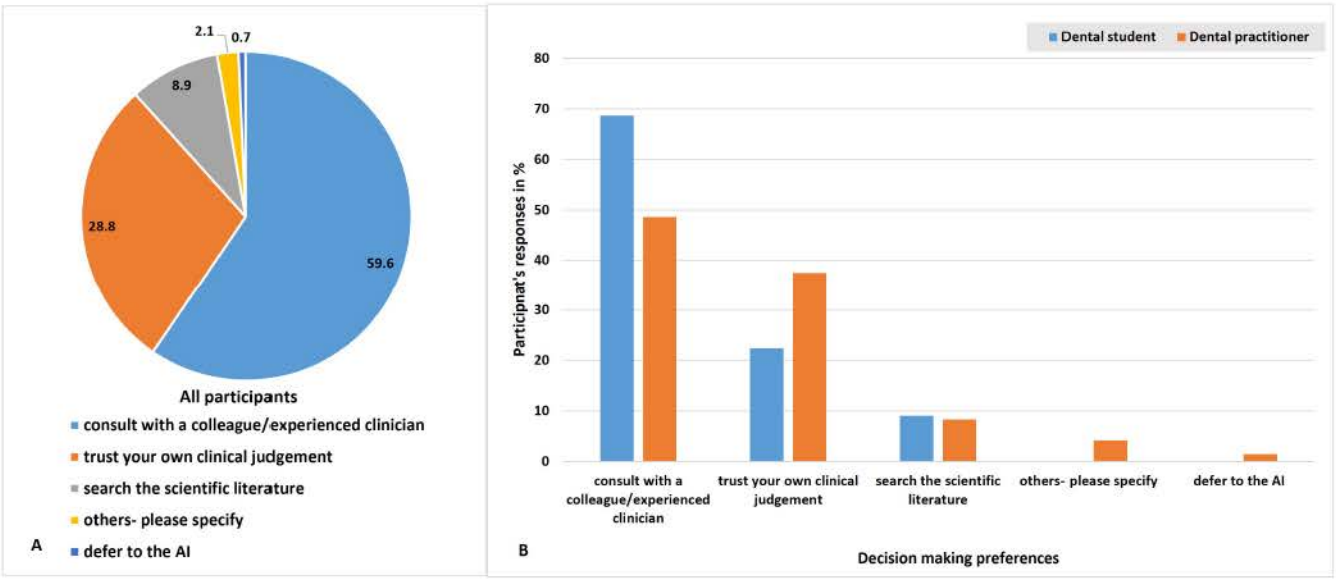


FIGURE 5 | Participant's preferences when their clinical decision differs from the AI decision. (A) All participants. (B) Participants' responses as a dental student or practitioner.

3.2 | Factors Affecting Attitudes Towards Dental AI Applications

The survey suggested that attitudes towards AI in dentistry were shaped by factors such as age, gender, clinical experience, professional qualifications and technological proficiency. Participants who were aware of dental AI applications were significantly older (mean age 35.4 ± 14.81 years) compared to the participants who were not aware of these applications (mean age 28.61 ± 7.35 years) ($p < 0.05$). Participants who believed AI's routine use was imminent within the next 5–10 years were older than participants who saw AI as already integrated or expected it to happen within 5 years. Gender differences were observed in the expectations for AI performance and its role in clinical support. More females viewed AI as a clinical tool ($n = 48$, 65.8%) and relied on their judgement over AI in discrepancies ($n = 24$, 32.9%) compared to males. Male participants were more inclined to use AI applications if available within the year ($M = 19$, 24%; $F = 11$, 14%).

Participants familiar with AI applications reported more clinical experience than those unfamiliar ($p = 0.02$) (Figure 6). Awareness of specific software was higher among specialist dentists compared to GPs ($p < 0.01$). In addition, specialists ($n = 6$, 54.5%) had a more positive attitude towards AI compared to general practitioners ($n = 11$, 16.2%), likely due to higher exposure to AI and its applications within their field ($p < 0.01$). Additionally, those who perceived themselves as technologically proficient reported higher awareness of AI's dental applications ($p < 0.01$). However, this self-assessment about technological proficiency did not significantly influence opinions on the timeline for AI integration in dentistry.

3.3 | Correlations Among Survey Variables

Significant correlations were observed between the responses received from the participants (Table S1, Data S1). A statistically significant moderate correlation ($r = 0.42$) was observed between the participant's awareness of a specific AI software and their likelihood of using AI if it became available in the next year. Participants' perceptions of the impact of AI on dental practice and workflows also correlated with the likelihood of using AI technologies if they became available within the year ($r = 0.53$, $p < 0.05$). A weak correlation between participants' perceptions of the timeline of AI integration and their technological proficiency and clinical experience was observed ($p < 0.05$). The negative correlation between the participant's age and their opinion about the timeline for AI integration into dental practice suggested that younger participants anticipated an earlier adoption of AI applications in dentistry.

4 | Discussion

This survey investigated the user experience and acceptance to understand the knowledge and perceptions of Australian dentists and dental and oral health students towards applications of AI technologies in dentistry. Although several studies [25, 29, 34–38] have reported the perceptions regarding AI applications, understanding user experience and acceptance is limited in medicine and dentistry, particularly in the Australian context. With an increase in the applications of AI technologies in our day-to-day lives, such as chatbots, AI smart assistants and autonomous vehicles, these technologies are becoming increasingly familiar. So, it is no surprise that a majority of our participants were familiar with AI applications. However, there was a notable gap in knowledge about specific dental AI software, likely due to the limited adoption of AI technologies in clinical practice. Regarding the reliance and trust in dental AI applications, the results indicated a cautious optimism among participants about integrating AI into their practices. While there is a general willingness to adopt AI technologies, hesitance persists, reflecting the need for more tangible demonstrations of the effectiveness of AI in clinical settings. Concerns about data protection and

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

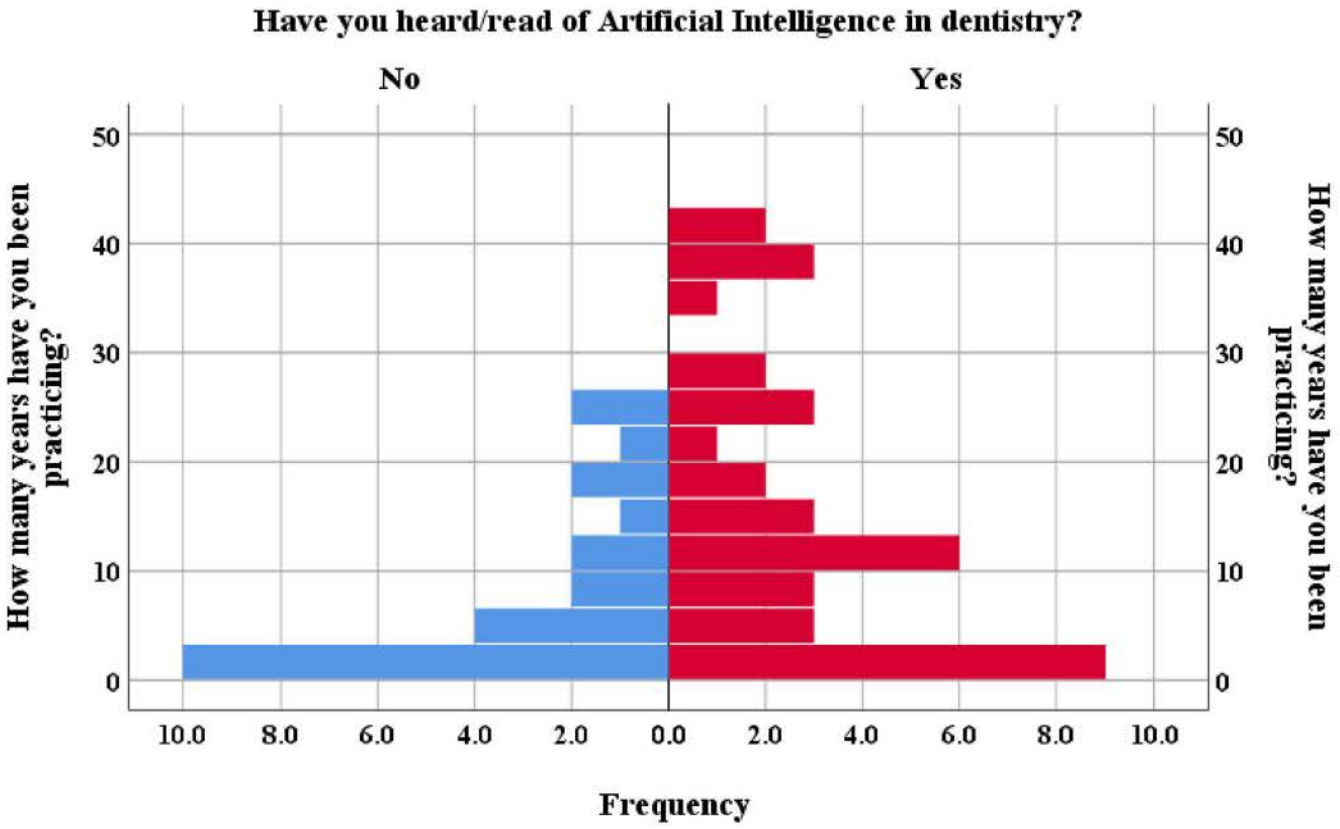


FIGURE 6 | Association between dentist's years in clinical practice and awareness of dental AI applications.

confidentiality [47] have already been identified, and governments worldwide have recognised this concern and created policies such as the EU's AI Act and Australia's Privacy Principles [48, 49]. Our findings are similar to those of other studies conducted in dental communities in other parts of the globe [33, 35, 46, 50, 51]. Across these studies, there is a generally positive attitude towards the potential benefits of AI in dentistry, such as improved diagnostic accuracy, efficiency and treatment planning. The knowledge regarding dental AI applications varied among dental students and practitioners in these studies. Lack of technical resources and training, concerns about data privacy, and algorithmic bias were identified as barriers to adopting AI technologies in these studies. The findings indicated a need for more focused educational efforts to enhance the awareness and understanding of dental AI applications as well as their capabilities and limitations.

The survey revealed that AI applications were perceived to have a broad scope for integration, particularly in dentomaxillofacial radiology and implantology. The specialisation of current AI algorithms in specific tasks such as caries detection limits its extensive use and integration in clinical practice [29]. However, with the fast-paced research and development in this area, we can expect the availability of more comprehensive AI applications that encompass all aspects of dentistry.

Our survey found that while there was considerable openness to adopting AI technologies in dental practice, there was some scepticism about AI decisions, where the participants preferred to consult a colleague or a senior clinician rather than rely on AI. It emphasises the need for AI technologies to be designed to augment rather than replace dentist's expertise. The participants' concerns about the impact of AI technologies on job security are valid and may stem from the potential for AI to automate certain aspects of dentistry [52]. However, it is essential to note that dentistry is a highly skilled profession and relies heavily on human expertise, empathy and nuanced judgement [53, 54]. There is a possibility that AI technologies will redefine job roles in dentistry, allowing dentists to focus on more complex cases or aspects of patient care. This misapprehension by dentists can be overcome by acquiring new skills related to AI and understanding how AI tools can enhance dental practice. The survey participants identified a lack of flexibility in the AI's decisions as a concern. The ability to make nuanced treatment decisions considering a patient's overall context, including quality of life, socioeconomic background, emotional well-being and personal preferences, is a uniquely human trait. These human experiences are complex, subtle and difficult to quantify and integrate into AI algorithms, making them incapable of incorporating the broader context of a patient's life into treatment decisions. For AI to be more widely adopted, robust regulatory and ethical guidelines are essential for safe clinical practice [55].

The differences in the perceptions between dentists and students underscore the importance of considering the role of professional experience in shaping perceptions of technological advancements. These findings reflect the findings in a recent systematic review conducted on the perceptions of dental students and practitioners regarding dental AI applications [56]. Dentists will likely view new technologies like AI through the lens of how these innovations fit into their practices. In contrast, students still in the formative stages of their careers may not have ingrained views

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ejc.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

or opinions. Understanding these differences is crucial for effectively communicating about and implementing AI technologies in dental practice and tailoring AI-related education and training for current and future dental professionals. It would be advantageous to conduct longitudinal studies to investigate the integration of AI into dental curricula and assess the effectiveness of different training methods. In addition, research into the impact of demographic factors (ethnic and cultural background of participants, clinician experience) and the long-term effects of AI integration in dentistry on diagnostic accuracy, treatment efficacy, patient outcomes and perceptions will provide deeper insight into the impact of AI technologies on clinical practice.

This study was limited by the sample size and the lack of representation of various demographic categories, including the cultural and ethnic backgrounds of participants. The voluntary nature of the survey may have resulted in selection bias, as only those participants with an interest in this topic may have responded to the survey. In addition, the self-reported data obtained from surveys are subject to response bias, the influence of the environment and context in which the survey was undertaken, leading to biased responses.

5 | Conclusion

The survey explored the user experience and acceptance to understand the current landscape of dental AI applications in Australia. The awareness of dental AI among dental students and practitioners is high. Our study suggests that the best use of dental AI applications would be as a support system that can provide data-driven insights, allowing dentists to focus on more patient-centred aspects of dental practice. Despite the concerns about the impact of AI on jobs and patient care, the participants foresee the integration of AI in dental care and professional practice in the near future. The study also recognised the various factors that affected the participant's attitudes towards AI. The opinion of the main stakeholders, including dental practitioners, students, education providers, policymakers and patients, is essential as it will significantly affect the integration of AI technologies in dentistry. Future strategies for AI implementation should consider ethical and regulatory challenges. Consequently, dental education and training programmes must adapt to include AI literacy, preparing dental practitioners to confidently and efficiently utilise AI technologies.

Acknowledgements

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. T. Shan, F. R. Tay, and L. Gu, "Application of Artificial Intelligence in Dentistry," *Journal of Dental Research* 100, no. 3 (2020): 232–244, https://doi.org/10.1177/0022034520969115.

2. V. Rajaraman, "JohnMcCarthy—Father of Artificial Intelligence," *Resonance* 19, no. 3 (2014): 198–207, https://doi.org/10.1007/s12045-014-0027-9.

3. A. Bohr and K. Memarzadeh, "The Rise of Artificial Intelligence in Healthcare Applications," in *Artificial Intelligence in Healthcare* (Cambridge, MA: Academic Press, 2020), 25–60.

4. I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science* 2, no. 3 (2021): 160, https://doi.org/10.1007/s42979-021-00592-x.

5. F. Schwendicke, W. Samek, and J. Krois, "Artificial Intelligence in Dentistry: Chances and Challenges," *Journal of Dental Research* 99, no. 7 (2020): 769–774, https://doi.org/10.1177/0022034520915714.

6. J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, "Diagnosis and Prediction of Periodontally Compromised Teeth Using a Deep Learning-Based Convolutional Neural Network Algorithm," *Journal of Periodontal & Implant Science* 48 (2018): 114–123.

7. M. S. Kavitha, A. Asano, A. Taguchi, T. Kurita, and M. Sanada, "Diagnosis of Osteoporosis From Dental Panoramic Radiographs Using the Support Vector Machine Method in a Computer-Aided System," *BMC Medical Imaging* 12, no. 1 (2012): 1–11, https://doi.org/10.1186/1471-2342-12-1.

8. A. Frydenlund, M. Eramian, and T. Daley, "Automated Classification of Four Types of Developmental Odontogenic Cysts," *Computerized Medical Imaging and Graphics* 38, no. 3 (2014): 151–162, https://doi.org/10.1016/j.compmedimag.2013.12.002.

9. K. Okada, S. Rysavy, A. Flores, and M. G. Linguraru, "Noninvasive Differential Diagnosis of Dental Periapical Lesions in Cone-Beam CT Scans," *Medical Physics* 42, no. 4 (2015): 1653–1665, https://doi.org/10.1118/1.4914418.

10. S. w. Chang, S. Abdul-Kareem, A. F. Merican, and R. B. Zain, "Oral Cancer Prognosis Based on Clinicopathologic and Genomic Markers Using a Hybrid of Feature Selection and Machine Learning Methods," *BMC Bioinformatics* 14, no. 1 (2013): 170, http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3673908&tool=pmcentrez&rendertype=abstract.

11. S. B. Khanagar, A. Al-Ehaideb, P. C. Maganur, et al., "Developments, Application, and Performance of Artificial Intelligence in Dentistry—A Systematic Review," *Journal of Dental Sciences* 16, no. 1 (2021): 508–522, https://www.sciencedirect.com/science/article/pii/S1991790220301434.

12. M. Johari, F. Esmaeili, A. Andalib, S. Garjani, and H. Saberkari, "Detection of Vertical Root Fractures in Intact and Endodontically Treated Premolar Teeth by Designing a Probabilistic Neural Network: An Ex Vivo Study," *Dento Maxillo Facial Radiology* 46, no. 2 (2017): 20160107.

13. M. Fukuda, K. Inamoto, N. Shibata, et al., "Evaluation of an Artificial Intelligence System for Detecting Vertical Root Fracture on Panoramic Radiography," *Oral Radiology* 36, no. 4 (2020): 337–343.

14. M. A. Saghiri, K. Asgar, K. K. Boukani, et al., "A New Approach for Locating the Minor Apical Foramen Using an Artificial Neural Network," *International Endodontic Journal* 45, no. 3 (2012): 257–265.

15. V. Patil, R. Vineetha, S. Vatsa, et al., "Artificial Neural Network for Gender Determination Using Mandibular Morphometric Parameters: A Comparative Retrospective Study," *Cogent Engineering* 7, no. 1 (2020): 1723783.

16. Pearl AI, "Pearl—The Future of Dentistry, Powered by Dental AI," (2022), accessed November 21, 2022, www.hellopearl.com.

17. Videa Health, "VideaHealth: The Dental AI Trusted by Dentists and DSOs," (2022), accessed November 21, 2022, https://www.videa.ai/.

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18. Denti AI, "Denti AI," (2022), accessed November 21, 2022, https://www.denti.ai/.

19. ORCA Dental AI, "ORCA Dental AI: Quality of Dental Care With the Power of AI," (2022), accessed November 21, 2022, https://www.orca-ai.com/.

20. Glidewell.Io In Office Solutions, "Glidewell.Io in Office Solutions," (2022), accessed November 21, 2022, https://glidewell.io/.

21. Smile Cloud Biometrics, "Smile Cloud Biometrics," (2022), accessed November 21, 2022, https://www.smilecloud.com/#intuitive.

22. Dental Monitoring, "Dental Monitoring," (2022), accessed November 21, 2022, https://dental-monitoring.com/home-3/.

23. V. Thorat, P. Rao, N. Joshi, P. Talreja, and A. R. Shetty, "Role of Artificial Intelligence (AI) in Patient Education and Communication in Dentistry," *Cureus* 16, no. 5 (2024): e59799.

24. T. Davenport and R. Kalakota, "The Potential for Artificial Intelligence in Healthcare," *Future Healthcare Journal* 6 (2019): 94–98.

25. M. Khalifa, M. Albadawy, and U. Iqbal, "Advancing Clinical Decision Support: The Role of Artificial Intelligence Across Six Domains," *Computer Methods and Programs in Biomedicine Update* 5 (2024): 100142, https://www.sciencedirect.com/science/article/pii/S2666990024000090.

26. A. Mahesh Batra and A. Reche, "A New Era of Dental Care: Harnessing Artificial Intelligence for Better Diagnosis and Treatment," *Cureus* 15, no. 11 (2023): e49319.

27. F. Pethani, "Promises and Perils of Artificial Intelligence in Dentistry," *Australian Dental Journal* 66, no. 2 (2021): 124–135, https://doi.org/10.1111/adj.12812.

28. H. Ding, J. Wu, W. Zhao, J. P. Matinlinna, M. F. Burrow, and J. K. H. Tsoi, "Artificial Intelligence in Dentistry—A Review," *Frontiers in Dental Medicine* 4 (2023): 1–13, https://www.frontiersin.org/articles/10.3389/fdmed.2023.1085251.

29. C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, "Key Challenges for Delivering Clinical Impact With Artificial Intelligence," *BMC Medicine* 17, no. 1 (2019): 195, https://doi.org/10.1186/s12916-019-1426-2.

30. H. Taherdoost, "A Review of Technology Acceptance and Adoption Models and Theories," *Procedia Manufacturing* 22 (2018): 960–967, https://www.sciencedirect.com/science/article/pii/S2351978918304335.

31. B. Rahimi, H. Nadri, H. Lotfnezhad Afshar, and T. Timpka, "A Systematic Review of the Technology Acceptance Model in Health Informatics," *Applied Clinical Informatics* 9, no. 3 (2018): 604–634.

32. V. Distler, C. Lallemand, and V. Koenig, "How Acceptable is This? How User Experience Factors can Broaden our Understanding of the Acceptance of Privacy Trade-Offs," *Computers in Human Behavior* 106 (2020): 106227, https://www.sciencedirect.com/science/article/pii/S0747563219304467.

33. R. Pauwels and Y. C. Del Rey, "Attitude of Brazilian Dentists and Dental Students Regarding the Future Role of Artificial Intelligence in Oral Radiology: A Multicenter Survey," *Dento Maxillo Facial Radiology* 50, no. 5 (2021): 20200461, https://doi.org/10.1259/dmfr.20200461.

34. M. Karan-Romero, R. E. Salazar-Gamarra, and X. A. Leon-Rios, "Evaluation of Attitudes and Perceptions in Students About the Use of Artificial Intelligence in Dentistry," *Dentistry Journal* 11, no. 5 (2023): 125, https://www.mdpi.com/2304-6767/11/5/125.

35. H. L. Abouzeid, S. Chaturvedi, K. M. Abdelaziz, F. A. Alzahrani, A. A. S. AlQarni, and N. M. Alqahtani, "Role of Robotics and Artificial Intelligence in Oral Health and Preventive Dentistry—Knowledge, Perception and Attitude of Dentists," *Oral Health & Preventive Dentistry* 19, no. 1 (2021): 353–363, http://www.ncbi.nlm.nih.gov/pubmed/34259428.

36. K. T. Aboalshamat, "Perception and Utilization of Artificial Intelligence (AI) Among Dental Professionals in Saudi Arabia," *Open Dentistry Journal* 16, no. 1 (2022): 1–7.

37. C. Sit, R. Srinivasan, A. Amlani, et al., "Attitudes and Perceptions of UK Medical Students Towards Artificial Intelligence and Radiology: A Multicentre Survey," *Insights Into Imaging* 11, no. 1 (2020): 14, https://doi.org/10.1186/s13244-019-0830-7.

38. J. Roganović, M. Radenković, and B. Miličić, "Responsible Use of Artificial Intelligence in Dentistry: Survey on Dentists' and Final-Year Undergraduates' Perspectives," *Health* 11, no. 10 (2023): 1480.

39. K. Cresswell, C. Sullivan, J. Theal, H. Mozaffar, and R. Williams, "Concerted Adoption as an Emerging Strategy for Digital Transformation of Healthcare—Lessons From Australia, Canada, and England," *Journal of the American Medical Informatics Association* 31 (2024): 1211–1215, https://doi.org/10.1093/jamia/ocae034.

40. Y. Chen, S. Khalid Khan, N. Shiwakoti, P. Stasinopoulos, and K. Aghabayk, "Analysis of Australian Public Acceptance of Fully Automated Vehicles by Extending Technology Acceptance Model," *Case Studies on Transport Policy* 14 (2023): 101072, https://www.sciencedirect.com/science/article/pii/S2213624X23001268.

41. F. Ullah, S. M. E. Sepasgozar, M. J. Thaheem, and F. Al-Turjman, "Barriers to the Digitalisation and Innovation of Australian Smart Real Estate: A Managerial Perspective on the Technology Non-adoption," *Environmental Technology & Innovation* 22 (2021): 101527, https://www.sciencedirect.com/science/article/pii/S2352186421001759.

42. L. Shinners, C. Aggar, S. Grace, and S. Smith, "Exploring Healthcare professionals' Understanding and Experiences of Artificial Intelligence Technology Use in the Delivery of Healthcare: An Integrative Review," *Health Informatics Journal* 26, no. 2 (2020): 1225–1236.

43. World Medical Association, "Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects," *Journal of the American Medical Association* 310, no. 20 (2013): 2191–2194.

44. L. A. Goodman, "Snowball Sampling," *Annals of Mathematical Statistics* 32, no. 1 (1961): 148–170, http://www.jstor.org/stable/2237615.

45. B. G. Glaser and A. L. Strauss, *Discovery of Grounded Theory: Strategies for Qualitative Research* (New York, NY: Transaction Publishers, 2017), https://www.frontiersin.org/articles/10.3389/fdmed.2023.1085251.

46. E. Yüzbaşıoğlu, "Attitudes and Perceptions of Dental Students Towards Artificial Intelligence," *Journal of Dental Education* 85, no. 1 (2021): 60–68, https://onlinelibrary.wiley.com/doi/abs/10.1002/jdd.12385.

47. M. A. Gianfrancesco, S. Tamang, J. Yazdany, and G. Schmajuk, "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data," *JAMA Internal Medicine* 178, no. 11 (2018): 1544–1547.

48. European Commission, "AI Act," 2024 Policies, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

49. Australian Government, "The Australian Privacy Principles," 2020 Office of the Australian Information Commissioner, https://www.oaic.gov.au/assets/privacy/australian-privacy-principles/the-australian-privacy-principles.pdf.

50. T. Eschert, F. Schwendicke, J. Krois, L. Bohner, S. Vinayahalingam, and M. Hanisch, "A Survey on the Use of Artificial Intelligence by Clinicians in Dentistry and Oral and Maxillofacial Surgery," *Medicina* 58, no. 8 (2022): 1059.

51. N. Singh, A. Pandey, A. P. Tikku, P. Verma, and B. P. Singh, "Attitude, Perception and Barriers of Dental Professionals Towards Artificial Intelligence," *Journal of Oral Biology and Craniofacial Research* 13, no. 5 (2023): 584–588, https://www.sciencedirect.com/science/article/pii/S2212426823000854.

52. M. Vodanović, M. Subašić, D. Milošević, and P. I. Savić, "Artificial Intelligence in Medicine and Dentistry," *Acta Stomatologica Croatica* 57, no. 1 (2023): 70–84.

53. R. L. Caplin, "Dentistry—Art or Science? Has the Clinical Freedom of the Dental Professional Been Undermined by Guidelines,

16000379, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Authoritative Guidance and Expert Opinion?," *British Dental Journal* 230, no. 6 (2021): 337–343.

54. D. Jeffrey, "Empathy, Sympathy and Compassion in Healthcare: Is There a Problem? Is There a Difference? Does It Matter?," *Journal of the Royal Society of Medicine* 109, no. 12 (2016): 446–452, https://doi.org/10.1177/0141076816680120.

55. C. Mennella, U. Maniscalco, G. De Pietro, and M. Esposito, "Ethical and Regulatory Challenges of AI Technologies in Healthcare: A Narrative Review," *Heliyon* 10, no. 4 (2024): e26297.

56. M. Dashti, J. Londono, S. Ghasemi, et al., "Attitudes, Knowledge, and Perceptions of Dentists and Dental Students Toward Artificial Intelligence: A Systematic Review," *Journal of Taibah University Medical Sciences* 19, no. 2 (2024): 327–337, https://www.sciencedirect.com/science/article/pii/S1658361223002755.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

16000579, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eje.13042 by Shwetha Hegde - University Of Sydney , Wiley Online Library on [16/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

	Technological proficiency	Are you currently aware of any AI software used in dentistry?	How far in the future do you think it will be before AI is integrated into dentistry?	Do you think that applications of AI in dentistry should be a part of the dental and oral health curriculum?	How do you see AI impacting your future workflow, tasks and/or practice of dentistry and oral health? AI will;	How do you think artificial intelligence will change the role of the clinician? AI will;	If AI were integrated into your clinic within the year, the likelihood of you using this technology is;	The idea of AI being integrated into your practice of dentistry and oral health;	At what level of proficiency do you expect the AI to perform?	If your clinical judgement and AI differ, would you; - Selected Choice
Participant's age		-0.287**	-0.231**							
Technological proficiency					-0.176*					
Are you currently aware of any AI software used in dentistry?			0.384**	0.273**	0.178*	0.175*	**0.418****	0.254*		
How far in the future do you think it will be before AI is integrated into dentistry?					0.278**	0.248**	0.280*	0.245*		
How do you see AI impacting your future workflow, tasks and practice of dentistry?						**0.425****	**0.530****	0.387**	-0.163*	
How do you think AI will change the role of the clinician?							0.387**	0.397**	-0.223**	
If AI were integrated into your clinic within the year, the likelihood of you using this technology is								**0.453****		
At what level of proficiency do										0.164*

you expect the AI to perform?										

Table S1: Correlation between various aspects of perception of dental AI technologies

* Indicates p- value <0.05; **indicates p-value<0.001

Supplementary File 1

Survey Flow

Block: Default Question Block (33 Questions)

Page Break

Start of Block: Default Question Block

Q1
Thank you for your interest in this study.

Please read the Participant Information Statement attached before continuing. By submitting your survey, you consent to take part in the study, and you are consenting to the use of your information. Once responses are submitted, they cannot be withdrawn as surveys are completed anonymously.

Participant information statement

○ I have read the participant information statement and I consent to participating in this study (1)

Q2 **Attitudes and Perceptions of Australian Dental Practitioners and Dental Students to Applications of Artificial Intelligence in Dentistry: A Survey**
Artificial Intelligence (AI) refers to a system which attempts to simulate the human cognition and critical thinking, through interconnected computer processes which may have the ability to self-learn.

The application of artificial intelligence to dentistry is an emerging area of interest. Current research focuses on where and how this developing technology can be applied.

Due to its novel nature, there is not yet an established consensus on how the dental community will welcome this technology.This survey aims to investigate the attitudes of dental practitioners and students on this topic. **disclaimer; for the purposes of this study we use 'AI' to encompass the terms and concepts of machine learning, deep learning, artificial intelligence, and algorithms.*

Page Break

Q3 Please select one option below
I am a;

- ○ Dental Student (1)
- ○ Oral Health Student (2)
- ○ Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist) (3)
- ○ Non-Registered Australian Dental Practitioner (4)
- ○ Other (5)

Skip To: End of Survey If Please select one option below I am a; = Non-Registered Australian Dental Practitioner

Skip To: End of Survey If Please select one option below I am a; = Other

Q4 Have you heard/read of AI (Artificial Intelligence) in dentistry?

- ○ Yes (1)
- ○ No (2)

Skip To: Q6 If Have you heard/read of AI (Artificial Intelligence) in dentistry? = No

Q5 Please select and option below for where you heard/read about it
You may select more than one option

- [ ] In clinical practice (1)
- [ ] Social media (Facebook, Twitter, LinkedIn etc.) (2)
- [ ] Media (Newspaper, News, Podcast etc.) (3)
- [ ] Lectures/Education/Dental Journals/Conferences (4)
- [ ] Patients (5)
- [ ] Other - please specify (6) _____

Q6 Are you currently aware of any AI software used in dentistry?

- ○ Yes (1)
- ○ No (2)

Skip To: Q8 If Are you currently aware of any AI software used in dentistry? = No

Q7 Please select how you were made aware of this technology
You may select more than one option

- [ ] In clinical practice (1)
- [ ] Social media (Facebook, Twitter, LinkedIn etc.) (2)
- [ ] Media (Newspaper, News, Podcast etc.) (3)
- [ ] Lectures/Education/Dental Journals/Conferences (4)
- [ ] Patients (5)
- [ ] Other - please specify (6) _____

Q8 How far in the future do you think it will be before AI is integrated into dentistry?

- ○ Already Integrated (1)
- ○ Within a year (2)
- ○ 1-5 years (3)
- ○ 5-10 years (4)
- ○ Never (5)
- ○ Not sure (6)

Q9 Do you think that applications of AI in dentistry should be a part of the dental and oral health curriculum?

- ○ Yes (1)
- ○ No (2)
- ○ Not sure (3)

Skip To: Q11 If Do you think that applications of AI in dentistry should be a part of the dental and oral health... = No

Skip To: Q12 If Do you think that applications of AI in dentistry should be a part of the dental and oral health... = Not sure

Q10 Please specify why you think that AI should be incorporated into the dental and oral health curriculum?

_____

Skip To: Q12 If Condition: Please specify why you thin... Is Displayed. Skip To: In which field do you think AI will b....

Q11 Please specify why you think AI should not be incorporated into the dental and oral health curriculum

_____

Q12 In which field do you think AI will be most applicable? You may select multiple options

- [ ] Dental Maxillofacial Radiology (1)
- [ ] Implantology (2)
- [ ] Restorative Dentistry (3)
- [ ] Endodontics (4)
- [ ] Oral Surgery (5)
- [ ] Orthodontics (6)
- [ ] Periodontics (7)
- [ ] Prosthodontics (8)
- [ ] Oral Pathology and Oral Medicine (9)
- [ ] Special Needs Dentistry (10)
- [ ] Paediatric Dentistry (11)
- [ ] Other - please specify (12) _____

Q13 In what task do you think AI would be most beneficial? - please click and drag the options from most to least beneficial

_____ Administrative Tasks - booking patients, taking payments (1)
_____ Image Processing - tooth numbering, anatomical structure delineation (2)
_____ Diagnosis - Analysis and localisation of pathology and presentation of differential diagnoses (3)
_____ Prognosis - prognostic tool to predict the course of disease and/or success of treatment (4)
_____ Treatment planning tool - risk evaluations, positioning of implants (5)

Q14 How do you see AI impacting your future workflow, tasks and/or practice of dentistry and oral health?
AI will;

- ○ Benefit all of dentistry (1)
- ○ Benefit some aspects of dentistry (2)
- ○ Make no difference to dentistry (3)
- ○ Negatively impact some aspects of dentistry (4)
- ○ Negatively impact all of dentistry (5)

Q15 How do you think artificial intelligence will change the role of the clinician?

AI will;

- ○ Assist and support the clinician (1)
- ○ Somewhat assist the clinician (2)
- ○ Have no impact on the clinician (3)
- ○ Somewhat negatively affect the clinician (4)
- ○ Negatively affect the clinician (5)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q16 If AI were integrated into your clinic within the year, the likelihood of you using this technology is;

- ○ Highly Likely (1)
- ○ Somewhat Likely (2)
- ○ Unchanged (3)
- ○ Somewhat Unlikely (4)
- ○ Highly Unlikely (5)

Display This Question:

If Please select one option below I am a; = Oral Health Student

Or Please select one option below I am a; = Dental Student

Q17 If AI were integrated into your education curriculum within the year, the likelihood of you applying this technology is;

- ○ Highly Likely (1)
- ○ Somewhat Likely (2)
- ○ Unchanged (3)
- ○ Somewhat Unlikely (4)
- ○ Highly Unlikely (5)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q18 The idea of AI being integrated into your practice of dentistry and oral health;

- Inspires excitement (1)
- Makes little difference (2)
- Raises concerns (3)

Skip To: Q21 If The idea of AI being integrated into your practice of dentistry and oral health; != Raises concerns

Display This Question:

If Please select one option below I am a; = Dental Student

Or Please select one option below I am a; = Oral Health Student

Q19 The idea of AI being integrated into your study of dentistry and oral health;

- Inspires excitement (1)
- Makes little difference (2)
- Raises Concerns (3)

Skip To: Q21 If The idea of AI being integrated into your study of dentistry and oral health; != Raises Concerns

Q20 The main concerns you would have are around;
You can select multiple options

- [ ] Security - patient confidentiality, risk of hackers (1)
- [ ] Redundancy - practitioners losing jobs to more efficient technology (2)
- [ ] Legality/Insurance - who is liable (3)
- [ ] Trust - inadequate information/expertise (4)
- [ ] Patient care - inflexibility of programme, basis on physiology/pathology alone (5)
- [ ] Accuracy - chances of AI being incorrect/making mistakes (6)
- [ ] Other - please specify (7) _____

Q21 At what level of proficiency do you expect the AI to perform?

- ○ Equivalent or superior to the worst performing specialist (1)
- ○ Equivalent to the average performing specialist (2)
- ○ Superior to the average performing specialist (3)
- ○ Equivalent to the best performing specialist (4)
- ○ Superior to the best performing specialist (5)

Q22 If your clinical judgement and AI differ, would you;

- ○ Defer to the AI (1)
- ○ Trust your own clinical judgement (2)
- ○ Consult with a colleague or another experienced clinician (3)
- ○ Perform literature research (4)
- ○ Other - please specify (5) _____

Page Break

Q23
Demographics Please specify your age

_____

Q24 Please indicate your gender

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

Display This Question:

If Please select one option below I am a; = Registered Australian Dental Practitioner (Dentist, Specialist, OHT, Hygienist)

Q25 Please select your dental field

- ○ General Practitioner (1)
- ○ Specialist (2)

Skip To: Q27 If Please select your dental field = General Practitioner

Skip To: Q29 If Please select your dental field = Specialist

Display This Question:

If Please select one option below I am a; = Dental Student

Or Please select one option below I am a; = Oral Health Student

Q26 Please select the type of degree and indicate which year of study you are currently in

- ○ Undergraduate - please indicate your current year of study (1)
  _____
- ○ Postgraduate - please indicate your current year of study (2)
  _____

Skip To: Q33 If Please select the type of degree and indicate which year of study you are currently in , Undergraduate - please indicate your current year of study Is Displayed

Q27 How many years have you been practicing?

_____

Q28 How many days do you work per week?

_____

Skip To: Q31 If Condition: How many days do you work p... Is Displayed . Skip To: Please select the options which descr....

Q29 Which field is your specialty?

_____

Q30 How many years have you been practicing?

_____

Q31 Please select the options which describe your current practice
You may select multiple options

- [ ] Metropolitan (1)
- [ ] Rural (2)
- [ ] Public (3)
- [ ] Private (4)
- [ ] Academia (5)

Q32 Please select all options which your current practice has in house
You may select multiple options

- [ ] OPG (1)
- [ ] CBCT (2)
- [ ] Digital Records (3)
- [ ] Digital Xray's (4)
- [ ] Cerec Machine (5)
- [ ] 3D scanner (6)

Q33 Please use your mouse to slide the bar to indicate how proficient you feel with technology in general

Unsatisfactory		Beginner		Satisfactory	Proficient
0	1	2	3	4	5

Technological Proficiency (1)	

End of Block: Default Question Block

Chapter 7

Machinometrics, a new framework towards explainable machine learning in healthcare.

Explainability of AI applications is essential as it ensures healthcare professionals understand and trust AI-driven decisions, leading to better patient outcomes. This chapter presents a paper that describes a novel method for providing explanations about AI outputs.

Title:

Machinometrics, a New Framework Towards Explainable Machine Learning in Healthcare

Under revisions BMC Methods

Author affiliations:

1. Mr Jionghui Lin (Co-First Author),
   Master of Data Science
   School of Computer Science,

   Faculty of Engineering,

   University of Sydney, Australia.

   Email: jlin0088@uni.sydney.edu.au

2. Dr Shwetha Hegde (Co-First Author),
   Dentomaxillofacial Radiology,

   Sydney Dental School,

   Faculty of Medicine and Health,

   University of Sydney, Australia.

   Email: shwetha.hegde@sydney.edu.au

3. Dr Daochang Liu (Co-First Author),
   School of Computer Science,

   Faculty of Engineering,

   University of Sydney, Australia.

   Email: Daochang.liu@sydney.edu.au

4. Mr Ka Wing Cheng
Master of Information Technology,

School of Computer Science,

Faculty of Engineering,

University of Sydney, Australia.

Email: phykawing@gmail.com

5. Prof Heiko Spallek,
Head of School and Dean,

Sydney Dental School,

Faculty of Medicine and Health,

University of Sydney, Australia.

Email: heiko.spallek@sydney.edu.au

6. A/Prof Chang Xu

School of Computer Science,

Faculty of Engineering,

University of Sydney, Australia.

Email: c.xu@sydney.edu.au

7. A/Prof Jinlong Gao,
Sydney Dental School,

Faculty of Medicine and Health,

University of Sydney, Australia.

Institute of Dental Research, Westmead Centre for Oral Health, NSW, Australia.

Jinlong.gao@sydney.edu.au

Corresponding authors:

Dr Daochang Liu,

J12- Computer Science Building,

University of Sydney

Daochang.liu@sydney.edu.au

A/Prof Jinlong Gao

Level 2, Westmead Centre for Oral Health C24, NSW 2145

PO Box 533, Wentworthville, NSW 2145, Australia

Jinlong.gao@sydney.edu.au

Abstract:

Background:

Machine learning (ML) offers significant potential in healthcare, but its clinical implementation is hampered by the 'black box nature' of these algorithms. This lack of transparency necessitates explainable models to ensure that the machine's decisions are interpretable and can be trusted by healthcare professionals. To address this issue, we introduce a novel Physically Grounded Explanation Framework using '*Machinometrics*' that aims to make the machine's decisions more understandable and relatable to healthcare professionals. We used the Machinometrics framework to enhance the interpretability of ML models in tooth identification tasks, a foundational application in clinical dentistry.

Methods:

The data set contained 1236 valid 3D CT scans of human teeth, categorised into eight classes. Data pre-processing converted DICOM files into NRRD format, isolating individual tooth volumes. The teeth data were split into five folds for cross-validation, and a 3D ResNet model was employed to identify the tooth. The model achieved an accuracy of 77.72%. Post-hoc analysis using Gradient-weighted Class Activation Mapping (Grad-CAM) was used to generate heatmap data, which were then correlated with the physical attributes such as density and distance to generate Machinometrics plots. Principal Component Analysis (PCA) was applied to further substantiate the model's decisions and reliability.

Results:

The case study on tooth classification demonstrated the potential of the Machinometrics framework in clinical applications. Density maps highlighted different tooth structures, such as enamel, dentine and pulp, while distance maps depicted the spatial relationship between tooth surfaces and surrounding space. PCA further substantiated the model's reliability, revealing significant patterns and correlations.

Discussion:

Machinometrics offers a promising avenue to recontextualise black-box models back to the physical world by serving as a common language among computer scientists and end users. By integrating physical attributes such as density and distance, we provided tangible, contextualised explanations of the model's decisions through 3D visualisations. This approach allows clinicians to intuitively assess and quantitatively evaluate the algorithm's outputs. This would, in turn, foster practical applications of advanced computational models in healthcare.

Keywords: Explainable Machine Learning, Healthcare Machine Learning, Physically Grounded Explanation

Introduction:

Machine Learning (ML) algorithms have been successfully incorporated across diverse spheres, including healthcare, finance, automotive and retail, demonstrating their versatility and widespread success in various applications.[1] In healthcare, ML applications have spanned different domains, including radiology, acute care, orthopaedics, cancer prediction and dentistry [2,3] and have been utilised for various clinical tasks, such as note-taking using speech recognition, clinical and radiographic image analysis using computer vision, drug monitoring, and object detection.[4,5] The availability of massive high-quality biological and clinical datasets and the development of different ML techniques have led to a surge in ML applications, as evidenced by the volume of scientific research and publications in this area.[6–10] ML applications have the potential to transform the delivery of healthcare by improving clinical efficiency, reducing diagnostic errors and making it more patient-centred and personalised.[7,11] However, this rapid and exponential growth in ML research is accompanied by challenges in translating research into real-world clinical applications. While the potential of ML to transform healthcare is undeniable, it is essential to address the inherent complexities of the algorithms. These challenges relate to ML model design, such as its black-box nature, algorithmic biases and brittleness.[12,13] Black-box models refer to algorithms whose internal workings are not transparent or understandable to humans. The lack of transparency poses risks such as lack of explainability, accountability,

responsibility, and trust, leading to challenges in deployment, especially in a high-stakes domain such as healthcare.[14,15] ML algorithms must have attributes such as explainability to assist human users effectively.[16] Explainable models are transparent models, and transparency connotes understanding how models work and how they might behave when presented with previously unseen data.[17] Explainability has emerged as an approach to assist clinicians and other healthcare workers in understanding the internal decisions, behaviour and actions of ML algorithms.[18] Explainability allows the users to be confident that the decisions generated by the ML algorithm are based on sound clinical guidelines and that the data underpinning these tools are accurate and reliable.[19]

Explainability can be instituted by making the model inherently explainable by design or by adding a second (*post hoc*) model to explain the black-box model.[20] Interpretable Fuzzy Systems, general linear models, and generalised additive models (GAM) are examples of inherently explainable models.[14] The *post hoc* explanations can be given as confidences, textual explanations, and visual explanations in the form of heat maps and pictorial representations.[21] Examples of the *post hoc* models include saliency-based methods such as attribution maps[22], heat maps[23] and Local Interpretable Model Agnostic Explanation (LIME).[24] The effectiveness of different forms of explanations has been explored, and one of the problems that has been recognised is that the explanations mainly use mathematical methods and are not directed at the end users, such as clinicians.

Instead, they are meant for computer scientists and developers.[14,25] Inherently explainable models are hindered by confounders, such as approximation errors, feature space complexity and the inability to detect and correct model errors.[26] The *post hoc* explainers can only approximate the mechanisms of the underlying ML algorithm to explain its decisions and, therefore, are not always accurate.[13] Therefore, prior post-hoc explanation methods fall short as they remain anchored in the model space[13] and rely on the human inspector to further implicitly make sense of the connection to the physical space, where ambiguity and biases are introduced. The saliency-based visual explanations are particularly useful in image recognition and classification tasks.[27] However, these explanations have intrinsic limitations since they are subject to human interpretation. There are concerns about oversimplification, misinterpretation, and fragility to adversarial attacks.[28] In addition, healthcare providers hold the duty of care and bear legal and moral responsibilities for clinical decisions. Given the current capabilities of ML algorithms, it is unlikely that machines will autonomously make clinical decisions in the near foreseeable future. The participation of the human user in this decision-making is critical, as a feasible and reasonable workflow integrating healthcare providers with ML algorithms can provide better patient outcomes. This calls for a more accessible and robust method of explanation that can be understood by healthcare providers and one that provides a common language about the inner workings of an ML model.

We argue that explaining an AI model is to bridge the inherent knowledge gap between models and humans, where the gap arises from the different modalities humans and models employ for decision-making. While humans receive information from the physical world and discern patterns based on tangible properties with physical meanings, models dwell in numerical tensors or matrices and try to decipher hidden patterns from the numbers, which are only indirect digital reflections of the physical world. In this sense, explainability involves forging a link between the model's decisions and the tangible world so that humans can validate the model's reasoning process against existing human knowledge rooted in physical reality. Hereby, we propose a new framework termed 'Physically Grounded Explanation' for black-box AI models, which builds a full chain from model input to model decisions by employing physical attributes extracted from input data to augment the explanation. This augmented explanation framework projects model explanation to physical representations, with an intuitive and relatable context for human users, enabling more straightforward auditing of model processes. In the healthcare domain, experts can explore the explanation using established human experience or domain knowledge to help build trust in AI models. Rather than confining explanations within the abstract space of the model, our proposed framework mitigates

ambiguity and subjectivity by explicitly mapping model decisions with real-world implications.

Concretely, as delineated in Figure 1, our proposed framework initiates the process by employing off-the-shelf *post-hoc* explanation techniques to generate a preliminary model explanation, subsequently augmented with physical attributes. The physical attributes are ideally computationally extracted from the input data, and its selection is guided by humans based on interest or domain knowledge. At the core of the framework is the concept of "*Machinometrics*", a term we introduce to describe the systematic integration between the model explanation and the physical attributes. Through this, we ground the explanation on tangible and understandable data elements. *Machinometrics* can be given to human users for qualitative and quantitative exploration to understand and scrutinise the model's decision process. In principle, the instantiation of the components in the framework is flexible; it can accommodate different explanation algorithms and physical attributes depending on the data modality, model architecture, and specific task need.

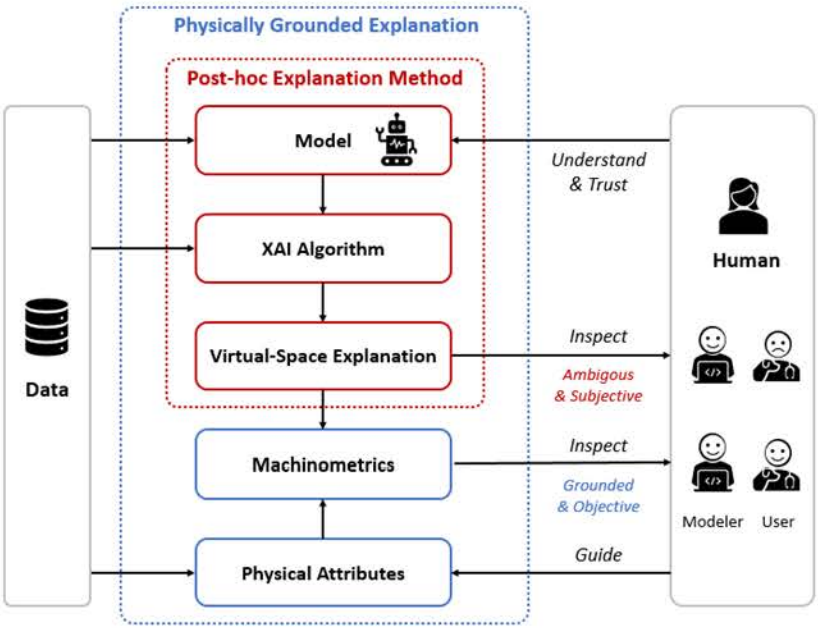


Figure 1: The schematic illustration of the proposed Physically Grounded Explanation framework. Red represents the existing methods and workflow to achieve explainability, and blue represents the proposed Physically Grounded Explanation framework.

Method:

We explore this novel physically grounded explanation framework to enhance the interpretability of machine learning (ML) models in tooth identification tasks. Tooth identification and classification is the foundation for ML applications in clinical dentistry. Convolutional neural networks (CNN) are widely used algorithms in the design of dental ML applications. *Post hoc* analyses have been used to provide visual explanations for the model's outputs.[29] However, as they do not provide enough context for why certain areas are highlighted, their interpretation can be ambiguous or too complex for clinicians.[30] In addition, the interpretation

of colour gradients and patterns can vary significantly among individuals, affecting the consistency and reliability of heat map interpretations.[31]

By situating the model's analytical outcomes within a three-dimensional visualisation, we provide an intuitive framework for clinicians to assess the model's decisions. To reflect how human users make decisions in tooth identification, physical attributes such as density and distance are integrated with heatmap data to generate the *Machinometrics* plot. Further advancing our methodology is using statistical methods such as Principal Component Analysis (PCA) to substantiate the model's decisions and reliability. We present a case study for applying the Physically Grounded Explanation framework using *Machinometrics* to explain the *post hoc* outputs by CNN models trained for tooth identification and classification from 3D CT images.

Data collection

Over 1,310 extracted human teeth were obtained from the disciplines of Tooth Morphology and Endodontics (Anonymised Dental School) and a dental student's private collection. The teeth were scanned using Computed Tomography (CT) (Canon Aquilion Lightning, Canon Medical systems), which was a 16-row, 32 slices helical CT scan with the following exposure and image acquisition settings: 80 kV, 30 mA, FOV 246, small focus, slice thickness 0.5, recon filter FC30 and paediatric scan 0.5 protocol. The original scanned CT volume was 512 x 512 x *L*, *L* is variable

depending on the length of the CT scan. There were 1,236 valid scanned teeth filed in our dataset for this study. Partial images lacking identifiable features were excluded from the dataset. The scanned teeth were labelled and categorised into 8 classes for this study by the academics from relevant disciplines at Anonymised Dental School. The teeth were categorised into 8 classes, including 250 upper incisors (UI), 87 lower incisors (LI), 184 upper canines (UC), 63 lower canines (LC), 296 upper premolars (UP), 94 lower premolars (LP), 155 upper molars (UM) and 107 lower molars (LM) (Extended Data File Figure 1). A unique accession number was assigned to each tooth collected in the dataset.

Data pre-processing

The CT data of collected human teeth were first retrieved in DICOM (Digital Imaging and Communications in Medicine) format. The DICOM data were then loaded by 3D Slicer and converted into NRRD (Nearly Raw Raster Data) format. The NRRD data can be easily read by Python using the pynrrd library as a numpy array. Since each scan had tens of extracted teeth arranged as an array lying on a low-density surface, each tooth was isolated first to get individual tooth volume.

The batch of scanned teeth was isolated from the background by the library connected components-3d (cc3d), with the threshold of minimum voxel intensity value of 500 and volume between 1,000 and 100,000 voxels. Each isolated volume of a tooth was exported as an individual NRRD file. The teeth dataset was randomly

split into five folds for stratified cross-validation, where the subsets preserve the ratio of each class.

Classification model

ResNet[39], a type of convolutional neural network (CNN) with residual connections, was selected as the machine learning algorithm towards teeth data. Since the teeth data was three-dimensional, 3D ResNet was utilised. A pre-trained model named r3d_18[40] was utilised. It was constructed with 18 layers based on ResNet. The basic block for each layer consists of a 3D convolution layer, batch normalisation and rectified linear unit (ReLU). This model was pre-trained with the Kinetics400 dataset[41], which was originally for the video classification task. The parameters in ResNet were fine-tuned on our teeth dataset, so that we could leverage prior knowledge learned from a larger dataset to improve performance on the teeth classification task. On the teeth data, the model was fine-tuned by minimising a negative log likelihood loss imposed on the predicted probability $p^{\hat{c}} \in [0,1]$ of the ground truth class $\hat{c}$ after the SoftMax function,

$$\mathcal{L} = -\log(p^{\hat{c}}).$$

The input image size was set to $\mathrm{H} \times \mathrm{W} \times \mathrm{D}$, which is $64 \times 64 \times 64$ in our case study. The overall average accuracy was 77.72% when we considered a five-fold cross-validation setting.

Training Details

The model was trained under the `PyTorch` framework Adam was utilised as the optimiser. The learning rate was set as 0.0001, while the batch size and epoch were set as 32 and 9, respectively. An NVIDIA A100 40 GB tensor core Graphics Processing Unit (GPU) was utilised for training.

Explainability algorithm

Gradient-weighted Class Activation Mapping (Grad-CAM)[42] was used as the post-hoc explainability algorithm to generate heatmaps indicating the vital regions of the image when the model made a prediction. Grad-CAM was chosen since it is suitable for a wide range of CNN-based architectures, including the ResNet used for our teeth classification. Grad-CAM utilises the gradient information in the CNN-based architecture to allocate importance weights to different neurons in an intermediate convolutional layer for a specific target class. Specifically, for a target c- th class, the attention heatmap map $A_m^c \in \mathbb{R}_{\geq 0}^{H' \times W' \times D'}$ with height H', width W', and depth D' of particular m-th sample is obtained via the following formulas. m denotes m -th sample in a particular c –th class. Height H', width W', and depth D'were set to be 8, 4 and 4, respectively, in our case study.

$$A_m^c = \mathrm{relu}(\textstyle\sum_k \alpha_k^c F^k),$$

$$\alpha_k^c = \frac{1}{Z} \textstyle\sum_h \sum_w \sum_d \frac{\partial y^c}{\partial F_{h,w,d}^k}.$$

The $F^k \in \mathbb{R}^{H' \times W' \times D'}$ represents the activations of k -th feature maps at the chosen convolutional layer, h, w, and d are indices for the spatial locations along the height, width, and depth dimensions, respectively, $y^c$is the predicted logit for class c before the SoftMax, ∂ represents the gradient operation and Z is a normalisation factor. The

attention heatmap is a linear combination of the feature maps weighted by the importance score α_k^c based on globally pooled gradients. After obtaining the attention map, it is scaled up to the original input image size $H \times W \times D$ through interpolation.

`Captum`, an extensible library for model explainability built on `PyTorch`, was utilised as a software package. Regarding our model, we chose the last convolutional layer when calculating GradCAM attribution. The training dataset was selected as the input for explainability analysis. Compared with the testing dataset, the training dataset can help us illustrate which parts of the images the model focused on during the training process. This can provide insights into how the model learns and which features it considers important for making predictions, which matches our study goal.

Physical attributes

Our proposed framework links the heatmap generated by the post-hoc explanation with physical attributes to provide a physically grounded explanation. Two physical attributes were calculated from the raw teeth data: density and distance. The raw intensity values within each tooth can be utilised to indicate the relative density information[43]. Thus, raw values within one tooth were normalised to act as normalised density values, which were selected as one of the physical attributes. The distance attribute was defined as the Euclidean distance of each foreground voxel (the voxel with a value higher than 0) within the tooth to the

nearest background voxel (the voxel with a value lower or equal to 0). Both attributes are normalised using the min-max scaling,

$$\bar{V} = (V - \min(V))/(\max(V) - \min(V)),$$

where $V \in \mathbb{R}_{\geq 0}^{H \times W \times D}$ and $\bar{V} \in [0,1]^{H \times W \times D}$ denote the attribute (density or distance) before and after normalisation, respectively.

***Machinometrics* plot integrating heatmap and physical attributes.**

To compute the physically grounded explanation, i.e., the correlation between the attention heatmap and the physical attribute, the range of the normalised physical attribute is first divided into N bins. For n-th bin, where $n \in \{0,1,\dots,N-1\}$, the range covers $\left(\frac{n}{N}, \frac{n+1}{N}\right]$. Then, we defined the correlation as a vector $C^m \in \mathbb{R}_{\geq 0}^{N}$ summarising the average attention value for voxels in each bin based on the physical attribute within the m -th sample.

$$C_n^m = \frac{1}{P}\frac{1}{Z'}\sum_h \sum_w \sum_d \mathbb{I}\left(\frac{n}{N} < \bar{V}_{h,w,d} \leq \frac{n+1}{N}\right) A^c_{m_{h,w,d}}$$

The $\mathbb{I}$ is an indicator function whose output is 1 when the condition holds and 0 otherwise, Z' is a normalisation factor and P is the number of voxels that within range from $\frac{n}{N}$ to $\frac{n+1}{N}$. Intuitively, C_n^m is the average heatmap value for all the voxels falling into the n -th bin defined by the physical attribute for the m -th sample. Since the correlation of each sample within its corresponding class may fluctuate, we defined the mean correlation vector calculated by each class as the final correlation vector $J^c \in \mathbb{R}_{\geq 0}^{N}$.

$$J_n^c = \frac{1}{Q}\sum_m C_n^m$$

Q represents the total number of samples in $c-$ th class. J_n^c is the average heatmap value for voxels in all samples in c-th class falling into the n-th bin. After combining the J_n^c with all bins as a vector, we can get the relationship between heatmap values and physical attributes, which can be taken as the plots in the Result Section we presented before. In the case study, we use 64 bins for density and 10 bins for distance.

Principal component analysis

We applied Principal Component Analysis (PCA) to heatmap values corresponding to the selected range of physical attributes derived from the *Machinometrics* plot. In PCA, each tooth sample can be taken as an observation, while its corresponding selected heatmap values are the variables. Since PCA is sensitive to the scale of the variables, we applied standardisation to variables across all observations before conducting PCA.

Results:

The results of the case study on tooth classification presented here underscore the potential of our approach in clinical applications to improve explainability and trust and make AI-driven diagnostics more accessible to clinicians and patients.

Integrating the *post hoc* analysis with physical attributes using *Machinometrics* generates plots for each class of tooth.

Density and distance were selected as physical attributes in *Machinometrics* analysis as they are relevant for tooth identification from CT scans. As illustrated in Figure 2.1.b and Figure 2.2.b, each scanned tooth comprises of regions of different densities, representing different structures of a tooth. The highest density region identified in the density map represents enamel, which is the outermost structure of the tooth crown. The lowest density region represents pulp tissue, which is the innermost structure of the tooth. The intermediate density region correlates with the dentine that is located between the enamel and pulp. As shown in Figure 2.1.c and Figure 2.2.c, the *post hoc* localisation map for each tooth was generated using Grad-CAM to visualise the importance measure of voxels used by the algorithm to make the decision. Using *Machinometrics* analysis, plots for each tooth in every class were generated by integrating density with heatmap data (d and e in Figures 2.1 and Figure 2.2). The plot of each tooth class was generated using the mean value of all plots in the relevant class, as presented in Figure 2.1.f and Figure 2.2.f.

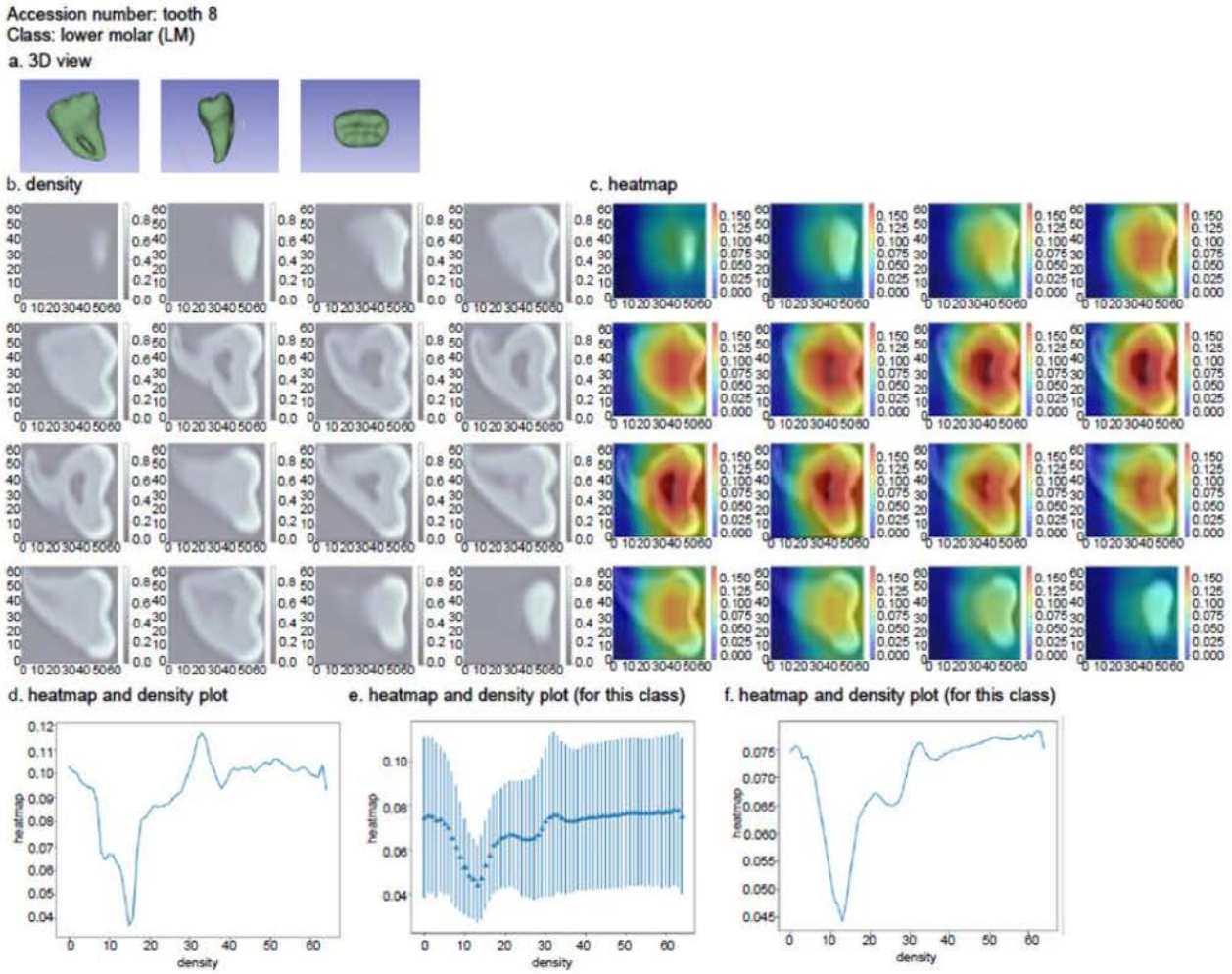


Figure 2.1: Representative *Machinometrics* plot by integrating density with heatmap for the nominated class of teeth. a. Buccal, proximal and occlusal views of 3D reconstructed CT images of selected tooth of the lower molar class (Accession number Tooth 8); b. Density maps of cross sections of tooth LM; c. Heatmaps for cross sections of tooth LM; d. Heatmap and density plot for Tooth 8; e. Heatmap and density plots for all teeth in LM; f. Mean heatmap and density plot for teeth in LM. LM, Lower Molar

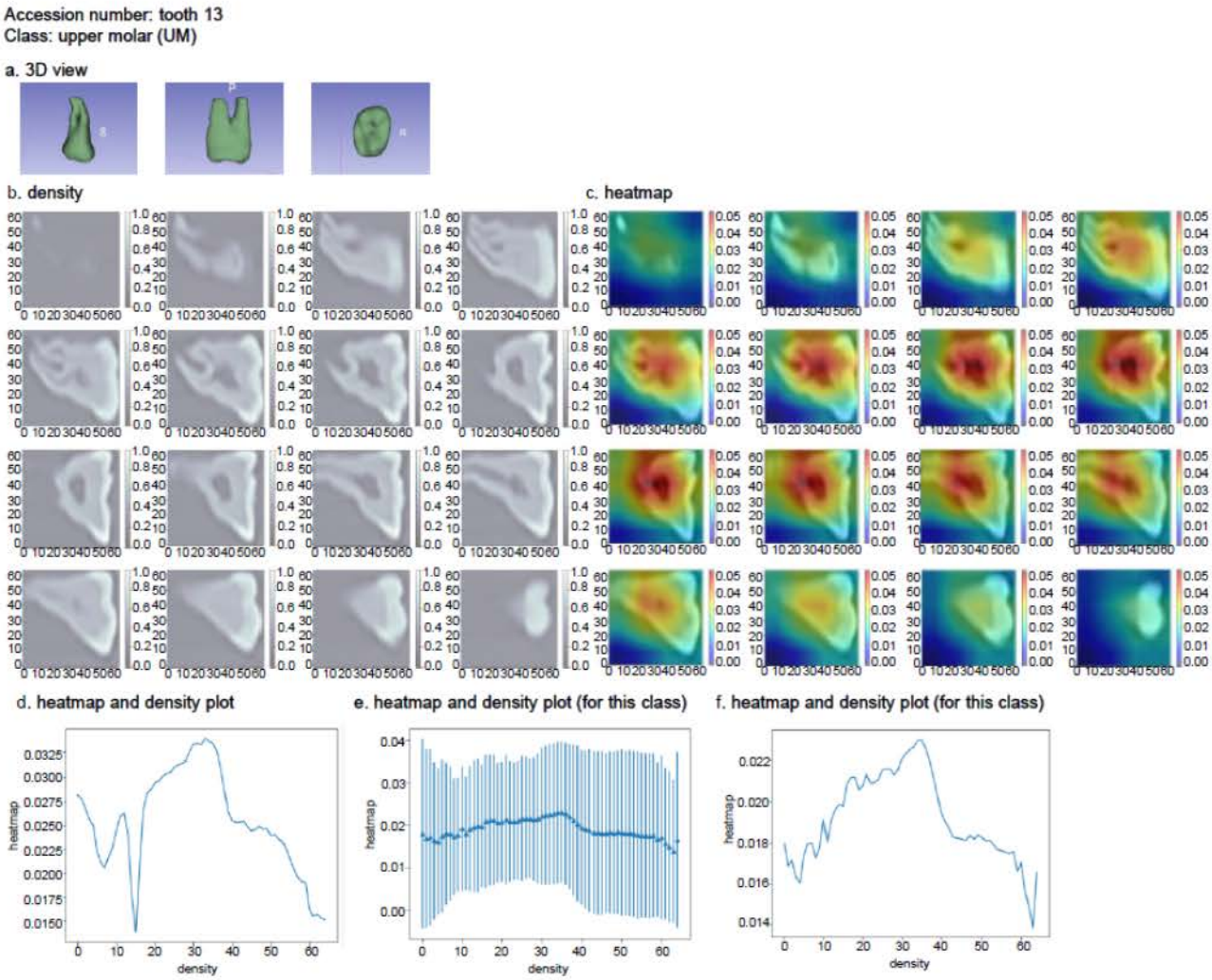


Figure 2.2: Representative Machinometrics plot by integrating density with heatmap for the nominated class of teeth. a. Buccal, proximal and occlusal views of 3D reconstructed CT images of selected tooth of the upper molar class (Accession number Tooth 13); b. Density maps of cross sections of tooth UM; c. Heatmaps for cross sections of tooth UM; d. Heatmap and density plot for Tooth 13; e. Heatmap and density plots for all teeth in UM class; f. Mean heatmap and density plot for teeth in UM class. UM, Upper Molar.

The second physical attribute selected in this study was distance, which is the shortest straight-line distance (Euclidean distance) between a voxel of the 3D tooth and the nearest background voxel. The background voxel refers to a voxel that is not a part of the tooth and is a part of the empty space in the 3D model. In

our study, this distance described the thickness of a tooth and determined the boundary between a tooth and the surrounding space. As shown in Figures 2.3.b and c as well as Figures 2.4.b and c, cross-sections and heat maps of representative LM and UM teeth were generated based on the distance between the surface of the teeth and the nearest background voxel. By integrating distance with heatmap data (Figure 2.3.d and Figure 2.4.d), plots for each tooth were generated to demonstrate the spatial information in the 3D image. The plot for tooth class LM and UM, together with the mean value of those two classes, are presented in Figures 2.3.e and f as well as Figures 2.4.e and f.

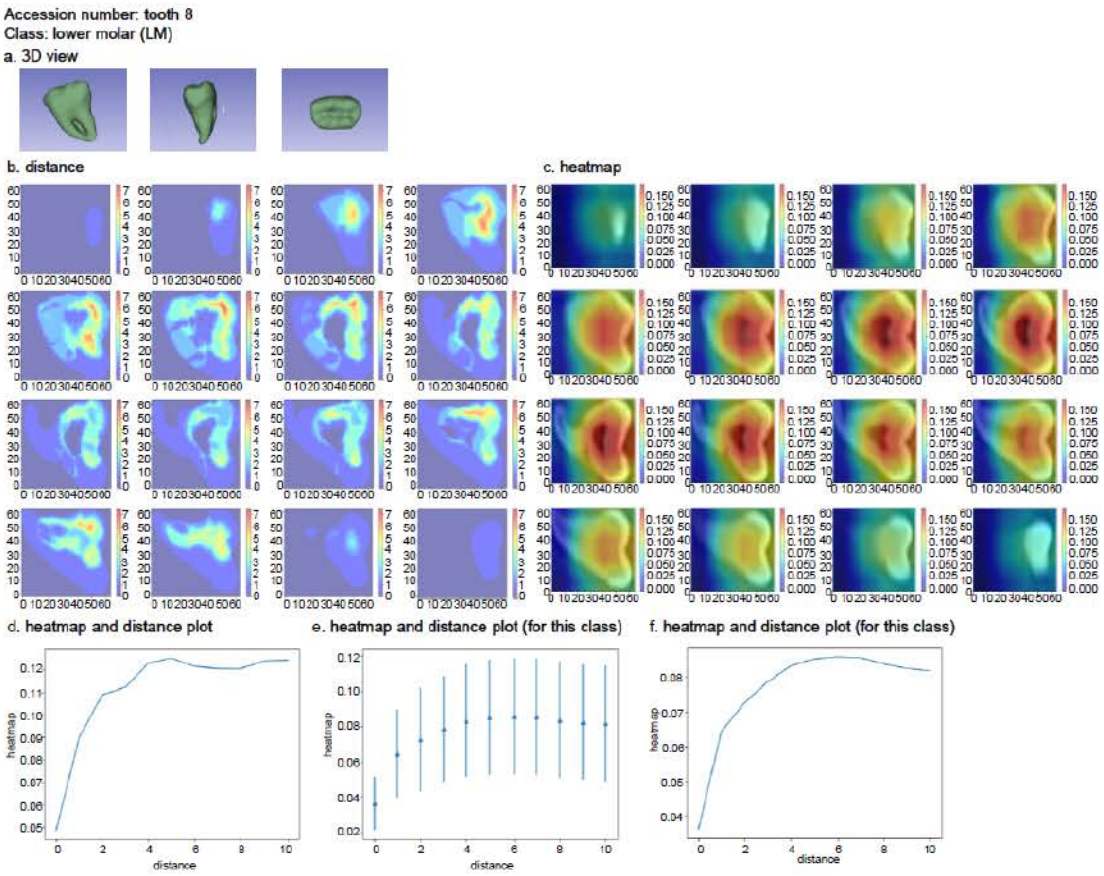


Figure 2.3: Representative *Machinometrics* plot by integrating distance with heatmap for the nominated class of teeth. a. Buccal, proximal and occlusal views of 3D reconstructed CT images of selected teeth for lower molar class (Accession number Tooth 8); b. Distance maps of cross sections of tooth LM; c. Heatmaps for cross sections of tooth LM ; d. Heatmap and distance plots for Tooth 8; e. Heatmap and distance plots for all teeth in LM class; f. Mean heatmap and distance plot for teeth in LM class. LM, Lower Molar.

The *Machinometrics* plots of the remaining 6 classes are presented in Extended Data File Figure 1.1 to Extended Data File Figure 2.6.

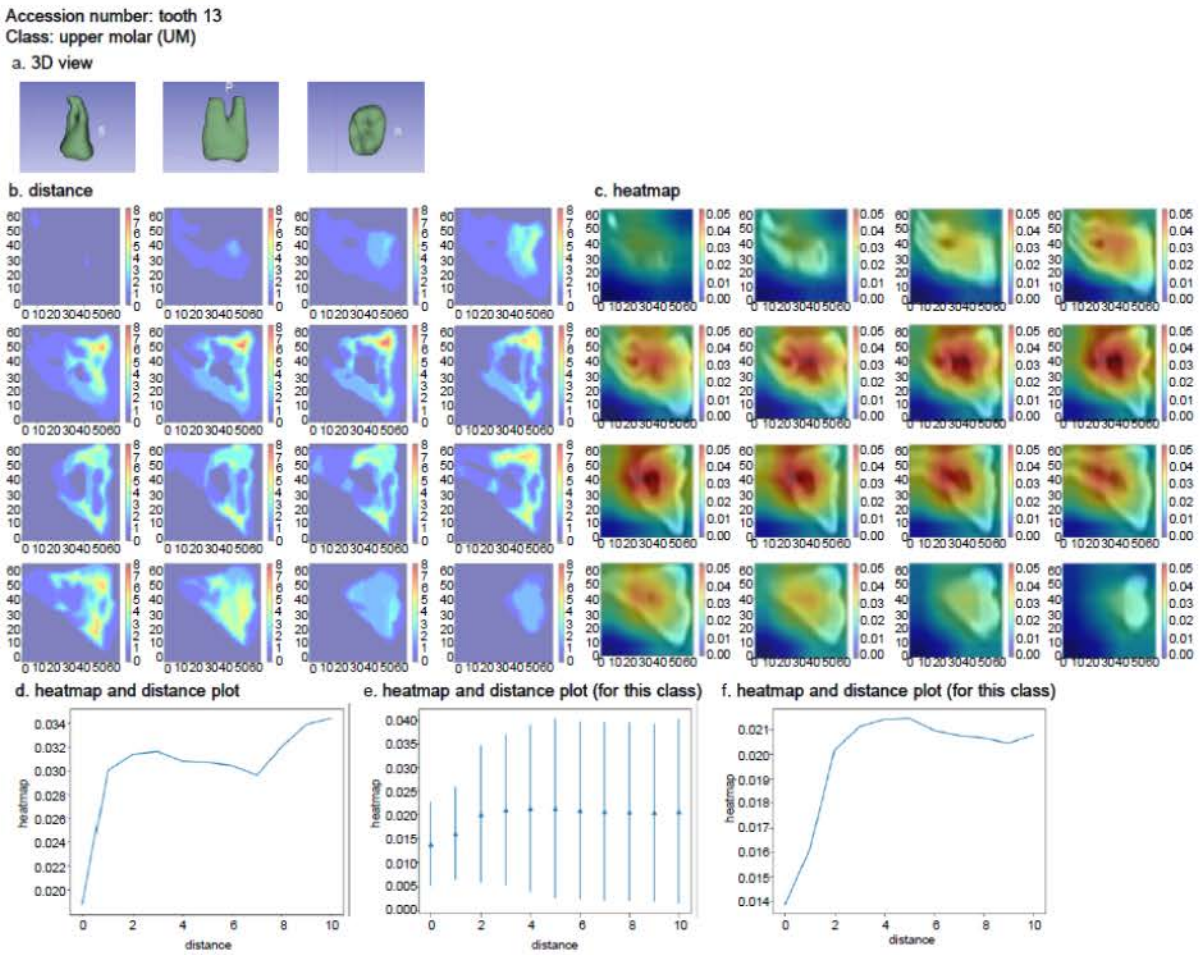


Figure 2.4: Representative *Machinometrics* plot by integrating distance with heatmap for the nominated class of teeth. a. Buccal, proximal and occlusal views of 3D reconstructed CT images of selected teeth for upper molar class (Accession number Tooth 13); b. Distance maps of cross sections of tooth UM; c. Heatmaps for cross sections of tooth UM; d. Heatmap and distance plots for Tooth 13; e. Heatmap and distance plots for all teeth in UM class; f. Mean heatmap and distance plot for teeth in UM class. UM, Upper Molar.

***Machinometrics* plots enable the interactive exploration of the algorithm's decisions by using visualisation and statistical analysis.**

Machinometrics plots created using the Physically Grounded Explanation Framework allowed human users and modellers to explore the machine's decisions interactively (Extended Data File 4: Video demo). Human users can select regions of interest in the *Machinometrics* plot to examine the decision by visualising in the physical world as well as evaluate the reliability of the decision using statistical tools such as Principal Component Analysis (PCA). When the explanations are grounded in physical attributes, the explanations become more meaningful to human users such as clinicians. Statistical analysis can reveal patterns and correlations in the data that might not be immediately apparent, helping to explain why the model makes certain decisions and classifications. PCA facilitates the visual interpretation of data by transforming it into principal components to reveal clusters, outliers or trends that correlate with the model's classification decisions.

A series of selected regions of interest from the plots of density and heatmap were chosen to reconstruct the 3D visualisation (Extended Data File Figure 3.1 to Extended Data File Figure 3.6). Dental academics in the research team then assessed whether the 3D visualisation represented the structural and morphological features of a tooth based on their domain knowledge. As

demonstrated in Figure 3.1, the density range between 30-35 for Tooth 8 and 31-36 for Tooth 13 in *Machinometrics* plots presented interpretable 3D visualisation resembling a tooth of LM and UM class, respectively. Using colours to represent different density ranges facilitated an intuitive understanding of relevant tooth structures. Interestingly, there was strong clustering of the density range between 30-35 in the PCA for the LM class, indicating this density range of interest could be used to explain the difference between the LM class and the other classes. Similarly, the PCA results suggested that the density range between 31-36 was a range of interest to identify the UM class from the other classes. Applying robust statistical confirmation enhanced the reliability and explainability of the model's decisions.

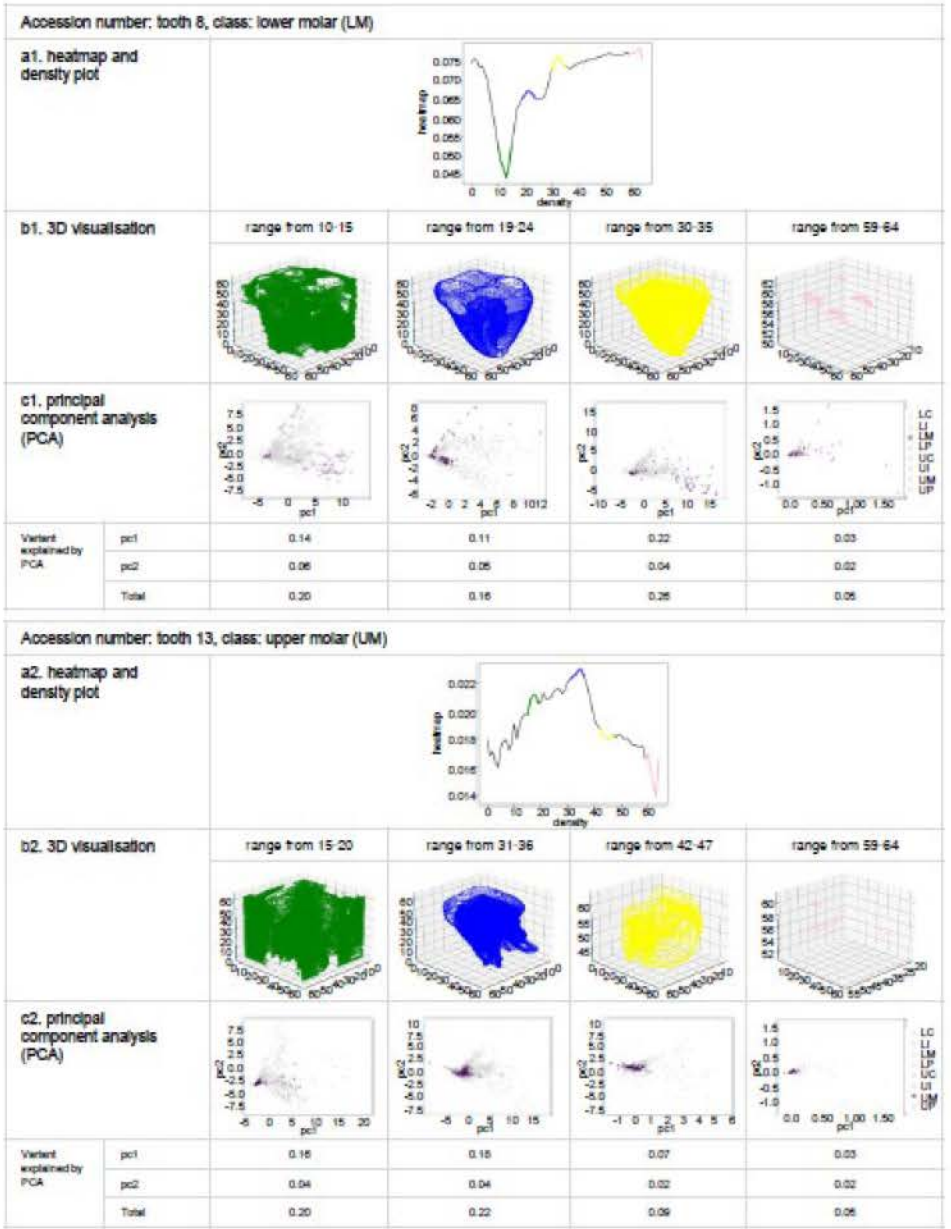


Figure 3.1: Representative 3D visualisation of density ranges representing different layers of teeth LM and UM with the Principal Component Analysis. a1 & a2. Heatmap and density plots for LM and UM (Accession numbers Tooth 8 and Tooth 13) with density ranges of user interest highlighted in different colours; b1 & b2. 3D visualisation of LM and UM in different ranges of density. c1 & c2. PCA shows clustering in different density ranges for LM, Lower Molar and UM, Upper Molar.

Further, similar analyses were conducted using the *Machinometrics* plot generated from the distance and heatmap data. As presented in Figure 3.2, the distance range between 0-2 in the *Machinometrics* plot for Tooth 13 in *Machinometrics* plots presented interpretable 3D visualisation resembling a tooth of LM and UM classes. Notably, this distance range represents the outermost surface of the tooth. The PCA results showed strong clustering for the shortest distance range to explain the differences among various tooth classes.

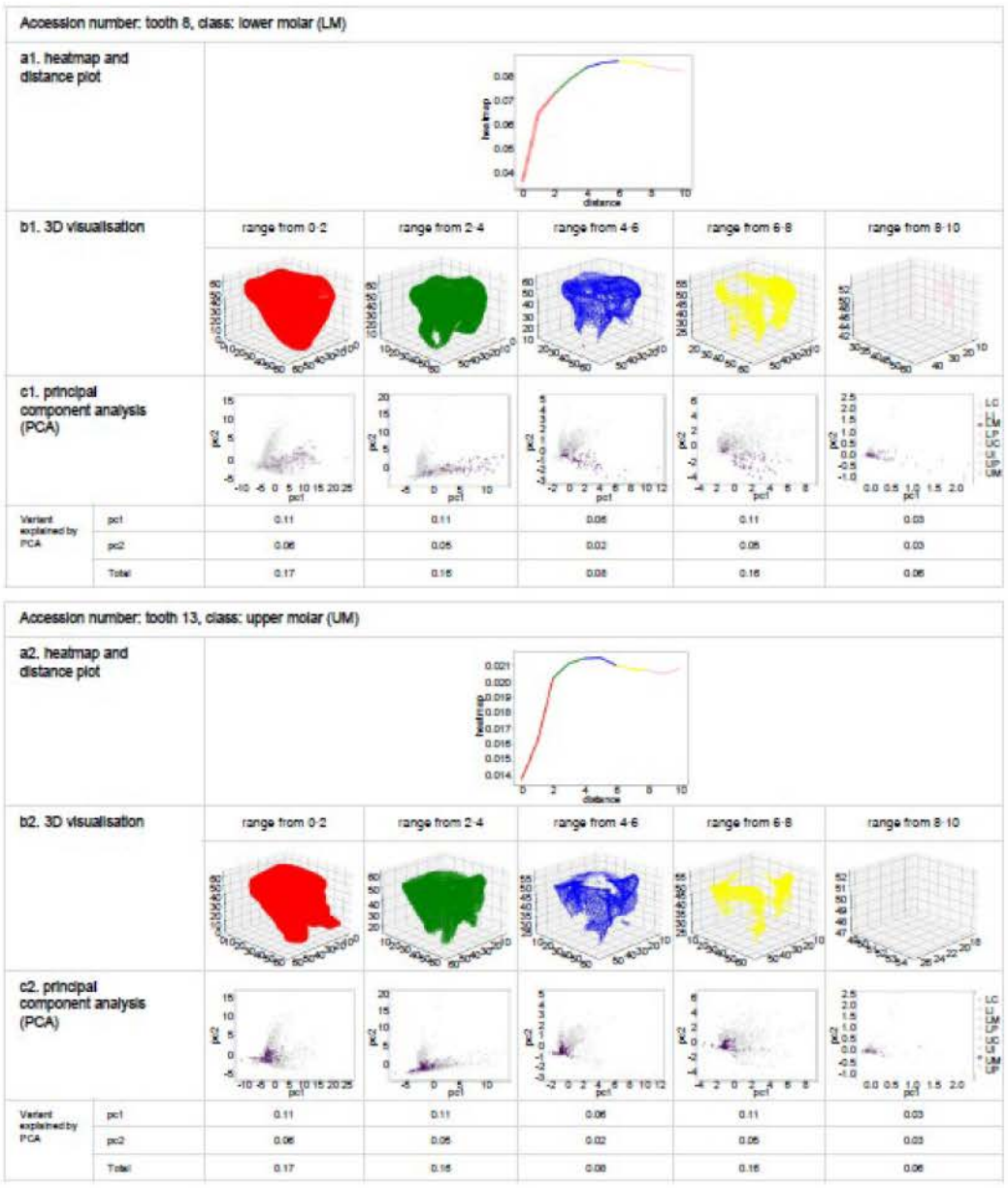


Figure 3.2: Representative 3D visualisation of distance ranges showing different layers of teeth LM and UM with Principal Component Analysis. a1 & a2. Heat map and distance plot for tooth LM and UM (Accession numbers Tooth 8 and tooth 13) with the distance ranges in different colours; b1 & b2. 3D visualisation of LM and UM in different distance ranges. c1 & c2. PCA clustering for different distance ranges of LM, lower molar and UM, upper molar.

Discussion:

ML has demonstrated great potential to improve diagnostic accuracy, treatment planning, and workflow efficiency in healthcare.[32] However, clinical implementation is limited due to the black-box nature of the algorithm design.[33] Currently, available *post hoc* models are developed to provide explanations to the modellers.[14,25] As a consequence, the interpretation of the ML algorithm's decisions is decontextualised. In addition, the *post hoc* models are limited by their abstract nature and the lack of interpretability to human users such as clinicians in healthcare.[30] It is pivotal to develop a 'common language' between the clinicians, computer scientists and ML models.

In this study, we developed the *Machinometrics* analysis in the Physically Grounded Explanation Framework, aiming to enhance the explainability of the model's decisions to human users. Human users develop their knowledge and understanding of physical reality. Introducing the physical attributes to *Machinometrics* analysis provides a tangible, contextualised and accessible approach to understanding the model's decisions. We demonstrated the application of *Machinometrics* analysis using a tooth identification task. By integrating physical attributes such as density and distance, we successfully extracted the information from the *Machinometrics* plot by generating 3D visualisations for human users to inspect the algorithm's decisions. Physical

attributes may vary depending on the different clinical contexts and the nature of the dataset. For example, density and distance were chosen in our case study as the algorithm was trained on CT images of teeth. *Machinometrics* plots also enabled human users to quantitively evaluate the algorithm's decisions using statistical analysis. The selection of the quantitative analysis method is also dependent on the context and the purpose determined by the human users.

Tooth identification and classification are fundamental parts of dentistry and are essential for ML algorithms used in dentistry. Correct identification and classification of teeth using ML methods can assist diagnosis and influence the choice of treatment methods, which can vary with the class of tooth and can lead to personalised patient care.[34] Accurate tooth identification also helps maintain detailed and precise patient records.[35] Tooth classification can aid in forensic investigations. Identifying teeth also plays a significant role in anthropology, as teeth are excellent indicators of age, dietary patterns, evolutionary changes, and dental disease history.[36] In addition, using ML models can serve as an educational tool to identify teeth and enhance training and education. Since the principles of image geometry and acquisition are similar in medicine and dentistry, these ML identification and classification methods demonstrated here are easily translatable to other areas in healthcare and beyond.

Our proposed Physically Grounded Explanation is a general and extendable framework. The components, including the model, the explanation algorithm, and the physical attributes, allow flexible instantiations. The framework is applicable to various tasks using a "black box" machine learning model, usually neural networks, where the architecture of the network is not constrained. The explanation algorithm is flexible and conditioned in that it provides a quantified explanation that can be matched to the input data to allow the association with the physical attributes. The explanation needs to be *post hoc* and compatible with the model used. As for the physical attribute, we define it in a broad sense as it should offer interpretable information linked with human knowledge in the real world. In practice, we suggest the physical attributes be identified by domain experts, and its implementation should be at a low cost, ideally automatically obtainable in a computational way.

NetDissection is an interpretability quantification method broadly related to our physically grounded explanation in spirit. NetDissection quantifies the interpretability of a neural network using the overlap between the activations of individual units and the segmentation masks of visual concepts.[38] Our framework differs in several ways. First, our framework provides a grounded explanation for users to explore and demystify the model decisions, which is different from the purpose of NetDissection to quantify the interpretability of a given model. Second, NetDissection measures the alignment between single units and visual concepts

such as objects, scenes, textures and colours annotated by humans. Our physically grounded explanation does not restrict the explanation at the unit level and allows flexibility depending on the choice of the post hoc explanation algorithm. The physical attribute also has a broader scope than the visual concept; it includes arbitrary characteristics in the physical space that can be either annotated or, ideally, computationally obtained. In this sense, the visual concepts in NetDissection are special cases of the physical attribute, and our physically grounded explanation is a generalised version of the NetDissection method.

Conclusion:

In conclusion, *Machinometrics* in the Physically Grounded Explanation Framework is an innovative approach to tackling the complex scenarios of ML design in healthcare. The future design of explainability in healthcare needs to recognise the complexity of healthcare, which involves various stakeholders, high-stakes decisions, patient safety, and economic and societal impacts. The current design of AI models follows a decontextualisation and reductionist paradigm, leading to conflict with complex scenarios in healthcare. Involving human users through the 'common language' in the design of explainability, machine learning models could be recontextualised for clinical applications.

List of abbreviations:

Machine learning (ML)

Generalised additive models (GAM)

Local interpretable model agnostic explanation (LIME)

Convolutional neural networks (CNN)

Principal component analysis (PCA)

Computed tomography (CT)

Upper incisors (UI)

Lower incisors (LI)

Upper canines (UC)

Lower canines (LC)

Upper premolars (UP)

Lower premolars (LP)

Upper molars (UM)

Lower molars (LM)

Digital imaging and communications in medicine (DICOM)

Nearly raw raster data (NRRD)

Rectified linear unit (ReLU)

Graphics processing unit (GPU)

Gradient-weighted class activation mapping (Grad-CAM)

Declarations:

Ethics approval and consent to participate

This project has received an exemption from ethics approval from the University of Sydney Human Ethics Office.

We have ensured that all data used in the study were handled according to the ethical guidelines and that the imaging data was stored securely in a password protected folder. The extracted teeth used in this project were collected over several years from various sources, including patients at Sydney Dental Hospital (SDH) and a private collection accumulated by a dental student. Patients at SDH sign a waiver permitting the use of extracted teeth for educational and research purposes. However, due to the anonymous nature and long period over which the teeth were collected, it is not feasible to obtain individual consent from each patient for the use of their teeth in this project. Additionally, the provenance of the teeth collected by the dental student is unknown. We believe that if the benefits of this research were explained to the patients, it is unlikely that they would have refused to consent to their deidentified images of teeth being used in this research.

Consent for publication

Our manuscript does not contain any individual person's identifiable data in any form (including individual details, images or videos). Therefore, consent for publication is not applicable.

Authors' contributions

J.L., S.H., and D.L. contributed equally to this work. J.L., S.H., D.L., K.W.C., C.X., J.G. conceived the method and designed the experiments. J.L. and K.W.C. performed the experiments. S.H. contributed materials and data. J.L., S.H., D.L., K.W.C., J.G. analysed the data. J.L., S.H., D.L., K.W.C., and J.G. wrote the manuscript. H.S. reviewed and revised the manuscript. All authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing interests.

Code and Data Availability

The code of the methodology and the dataset of 3D tooth CT images can be available upon request.

Acknowledgements:

We acknowledge Mr Peter O'Reilly, Lecturer in the Discipline of Medical Imaging Science, University of Sydney, for his invaluable expertise and assistance in scanning the natural teeth used in this research.

References:

1. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci [Internet]. 2021;2(3):160. Available from: https://doi.org/10.1007/s42979-021-00592-x

2. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ [Internet]. 2023;23(1):689. Available from: https://doi.org/10.1186/s12909-023-04698-z

3. Carrillo-Perez F, Pecho OE, Morales JC, Paravina RD, Della Bona A, Ghinea R, et al. Applications of artificial intelligence in dentistry: A comprehensive review. J Esthet Restor Dent [Internet]. 2022;34(1):259–80. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/jerd.12844

4. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. Nat Biomed Eng [Internet]. 2022;6(12):1330–45. Available from: https://doi.org/10.1038/s41551-022-00898-y

5. Shamshirband S, Fathi M, Dehzangi A, Chronopoulos AT, Alinejad-Rokny H. A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues. J Biomed Inform [Internet]. 2021;113:103627. Available from: https://www.sciencedirect.com/science/article/pii/S1532046420302550

6. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. J Dent Res [Internet]. 2020 Apr 21;99(7):769–74. Available from: https://doi.org/10.1177/0022034520915714

7. Hung K, Montalvo C, Tanaka R, Kawai T, Bornstein M. The use and performance of artificial intelligence applications in dentomaxillofacial radiology: a systematic review. Dentomaxillofacial Radiol. 2020;49(1).

8. Rubinger L, Gazendam A, Ekhtiari S, Bhandari M. Machine learning and artificial intelligence in research and healthcare. Injury [Internet]. 2023;54:S69–73. Available from: https://www.sciencedirect.com/science/article/pii/S0020138322000766

9. Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. Multimed Tools Appl [Internet]. 2023;82(17):26731–69. Available from: https://doi.org/10.1007/s11042-022-14305-w

10. Shehab M, Abualigah L, Shambour Q, Abu-Hashem MA, Shambour MKY, Alsalibi AI, et al. Machine learning in medical applications: A review of state-of-the-art methods. Comput Biol Med [Internet]. 2022;145:105458. Available from:

https://www.sciencedirect.com/science/article/pii/S0010482522002505

11. Pethani F. Promises and perils of artificial intelligence in dentistry. Aust Dent J [Internet]. 2021 Jun 1;66(2):124–35. Available from: https://doi.org/10.1111/adj.12812

12. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med [Internet]. 2019;17(1):195. Available from: https://doi.org/10.1186/s12916-019-1426-2

13. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell [Internet]. 2019;1(5):206–15. Available from: https://doi.org/10.1038/s42256-019-0048-x

14. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Inf Fusion [Internet]. 2023;99:101805. Available from: https://www.sciencedirect.com/science/article/pii/S1566253523001148

15. Kiseleva A, Kotzinos D, De Hert P. Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations. Front Artif Intell [Internet]. 2022;5. Available from: https://www.frontiersin.org/articles/10.3389/frai.2022.879603

16. Joyce DW, Kormilitzin A, Smith KA, Cipriani A. Explainable artificial intelligence for mental health through transparency and interpretability for understandability. npj Digit Med [Internet]. 2023;6(1):6. Available from: https://doi.org/10.1038/s41746-023-00751-9

17. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, et al. Explainable AI (XAI): Core Ideas, Techniques, and Solutions. ACM Comput Surv [Internet]. 2023 Jan;55(9). Available from: https://doi.org/10.1145/3561048

18. Yang CC. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. J Healthc Informatics Res [Internet]. 2022;6(2):228–39. Available from: https://doi.org/10.1007/s41666-022-00114-1

19. Simkute A, Luger E, Jones B, Evans M, Jones R. Explainability for experts: A design framework for making algorithms supporting expert decisions more explainable. J Responsible Technol [Internet]. 2021;7–8:100017. Available from: https://www.sciencedirect.com/science/article/pii/S266665962100010X

20. Hassija V, Chamola V, Mahapatra A, Singal A, Goel D, Huang K, et al. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. Cognit Comput [Internet]. 2023; Available from: https://doi.org/10.1007/s12559-023-10179-8

21. Vale D, El-Sharif A, Ali M. Explainable artificial intelligence (XAI) post-hoc explainability methods: risks and limitations in non-discrimination law. AI Ethics [Internet]. 2022;2(4):815–26. Available from: https://doi.org/10.1007/s43681-022-00142-y

22. Ma J, Schneider L, Lapuschkin S, Achtibat R, Duchrau M, Krois J, et al. Towards Trustworthy AI in Dentistry. J Dent Res [Internet]. 2022;101(11):1263–8. Available from: https://doi.org/10.1177/00220345221106086

23. Yoon K, Kim JY, Kim SJ, Huh JK, Kim JW, Choi J. Explainable deep learning-based clinical decision support engine for MRI-based automated diagnosis of temporomandibular joint anterior disk displacement. Comput Methods Programs Biomed. 2023;233:107465.

24. Bhattacharjee N. Automated Dental Cavity Detection System Using Deep Learning and Explainable AI. AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci. 2022;2022:140–8.

25. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020 Nov;20(1):310.

26. Poursabzi-Sangdeh F, Goldstein DG, Hofman JM, Wortman Vaughan JW, Wallach H. Manipulating and Measuring Model Interpretability. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems [Internet]. New York, NY, USA: Association for Computing Machinery; 2021. (CHI '21). Available from: https://doi.org/10.1145/3411764.3445315

27. Borys K, Schmitt YA, Nauta M, Seifert C, Krämer N, Friedrich CM, et al. Explainable AI in medical imaging: An overview for clinical practitioners – Beyond saliency-based XAI approaches. Eur J Radiol [Internet]. 2023;162:110786. Available from: https://www.sciencedirect.com/science/article/pii/S0720048X23001006

28. Ghorbani A, Abid A, Zou J. Interpretation of neural networks is fragile. In: Proceedings of the AAAI conference on artificial intelligence. 2019. p. 3681–8.

29. Turbé H, Bjelogrlic M, Lovis C, Mengaldo G. Evaluation of post-hoc interpretability methods in time-series classification. Nat Mach Intell [Internet]. 2023;5(3):250–60. Available from: https://doi.org/10.1038/s42256-023-00620-w

30. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform [Internet]. 2021;113:103655. Available from: https://www.sciencedirect.com/science/article/pii/S1532046420302835

31. Crameri F, Shephard GE, Heron PJ. The misuse of colour in science communication. Nat Commun [Internet]. 2020;11(1):5444. Available from: https://doi.org/10.1038/s41467-020-19160-7

32. Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database [Internet]. 2020 Jan 1;2020:baaa010. Available from: https://doi.org/10.1093/database/baaa010

33. Poon AIF, Sung JJY. Opening the black box of AI-Medicine. J Gastroenterol Hepatol

[Internet]. 2021 Mar 1;36(3):581–4. Available from: https://doi.org/10.1111/jgh.15384

34. Sadr S, Rokhshad R, Daghighi Y, Golkar M, Tolooie Kheybari F, Gorjinejad F, et al. Deep learning for tooth identification and numbering on dental radiography: a systematic review and meta-analysis. Dentomaxillofac Radiol. 2024 Jan;53(1):5–21.

35. Astekar M, Saawarn S, Ramesh G, Saawarn N. Maintaining dental records: Are we ready for forensic needs? J Forensic Dent Sci. 2011 Jul;3(2):52–7.

36. Jayakrishnan JM, Reddy J, Vinod Kumar RB. Role of forensic odontology and anthropology in the identification of human remains. J Oral Maxillofac Pathol. 2021;25(3):543–7.

37. Vandenberghe B. The digital patient – Imaging science in dentistry. J Dent [Internet]. 2018;74:S21–6. Available from: https://www.sciencedirect.com/science/article/pii/S0300571218301003

38. Bau D, Zhou B, Khosla A, Oliva A, Torralba A. Network Dissection: Quantifying Interpretability of Deep Visual Representations. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. p. 3319–27.

39. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. p. 770–8.

40. Tran D, Wang H, Torresani L, Ray J, LeCun Y, Paluri M. A Closer Look at Spatiotemporal Convolutions for Action Recognition. arXiv e-prints. 2017 Nov;arXiv:1711.11248.

41. Kay W, Carreira J, Simonyan K, Zhang B, Hillier C, Vijayanarasimhan S, et al. The Kinetics Human Action Video Dataset. arXiv e-prints. 2017 May;arXiv:1705.06950.

42. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Int J Comput Vis [Internet]. 2020;128(2):336–59. Available from: https://doi.org/10.1007/s11263-019-01228-7

43. Seeram E. Computed Tomography - E-Book: Computed Tomography - E-Book [Internet]. 5th edition. Elsevier Health Sciences; 2022. 49–63, 102–115 p. Available from: https://books.google.com.au/books?id=faZ1EAAAQBAJ

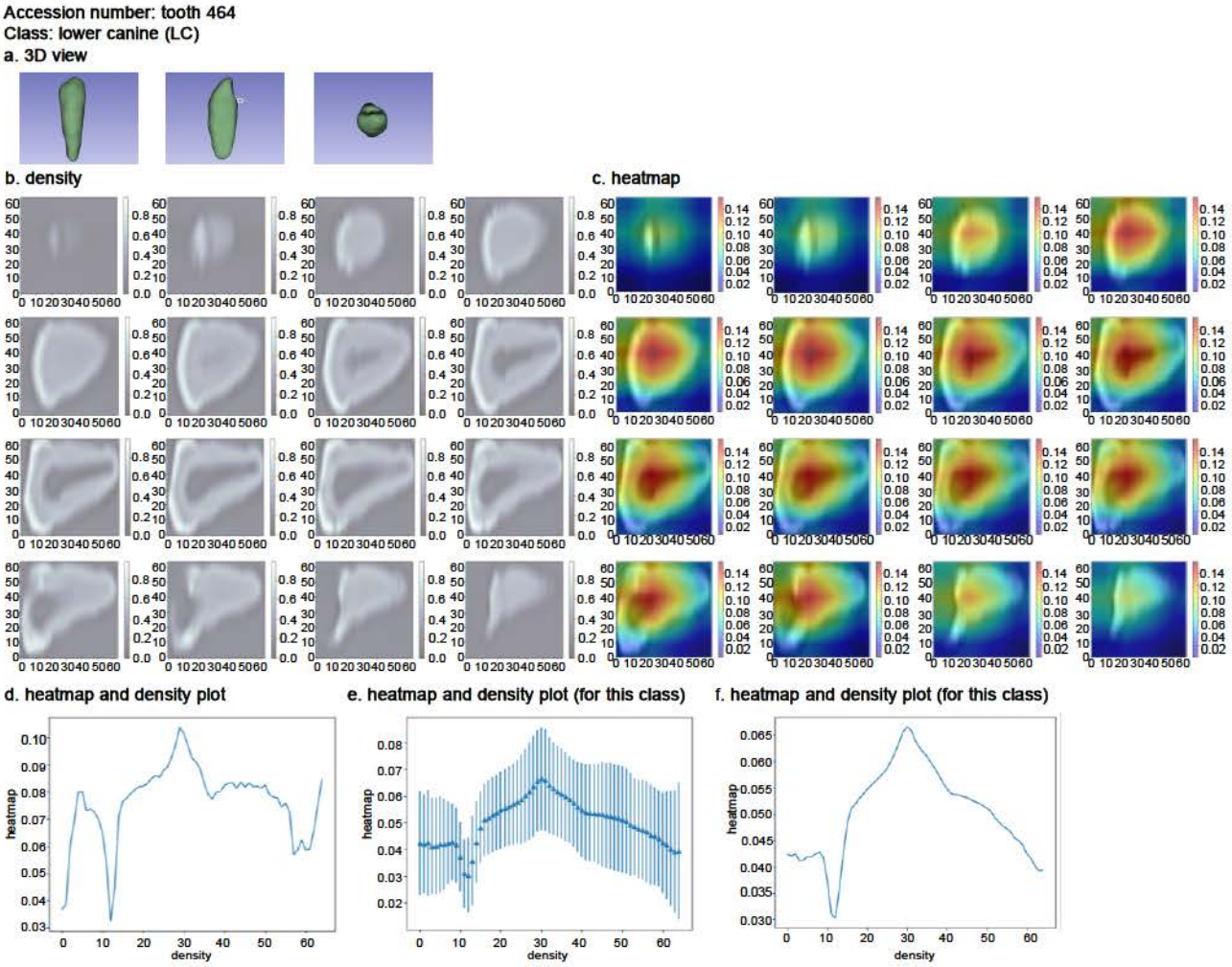
1



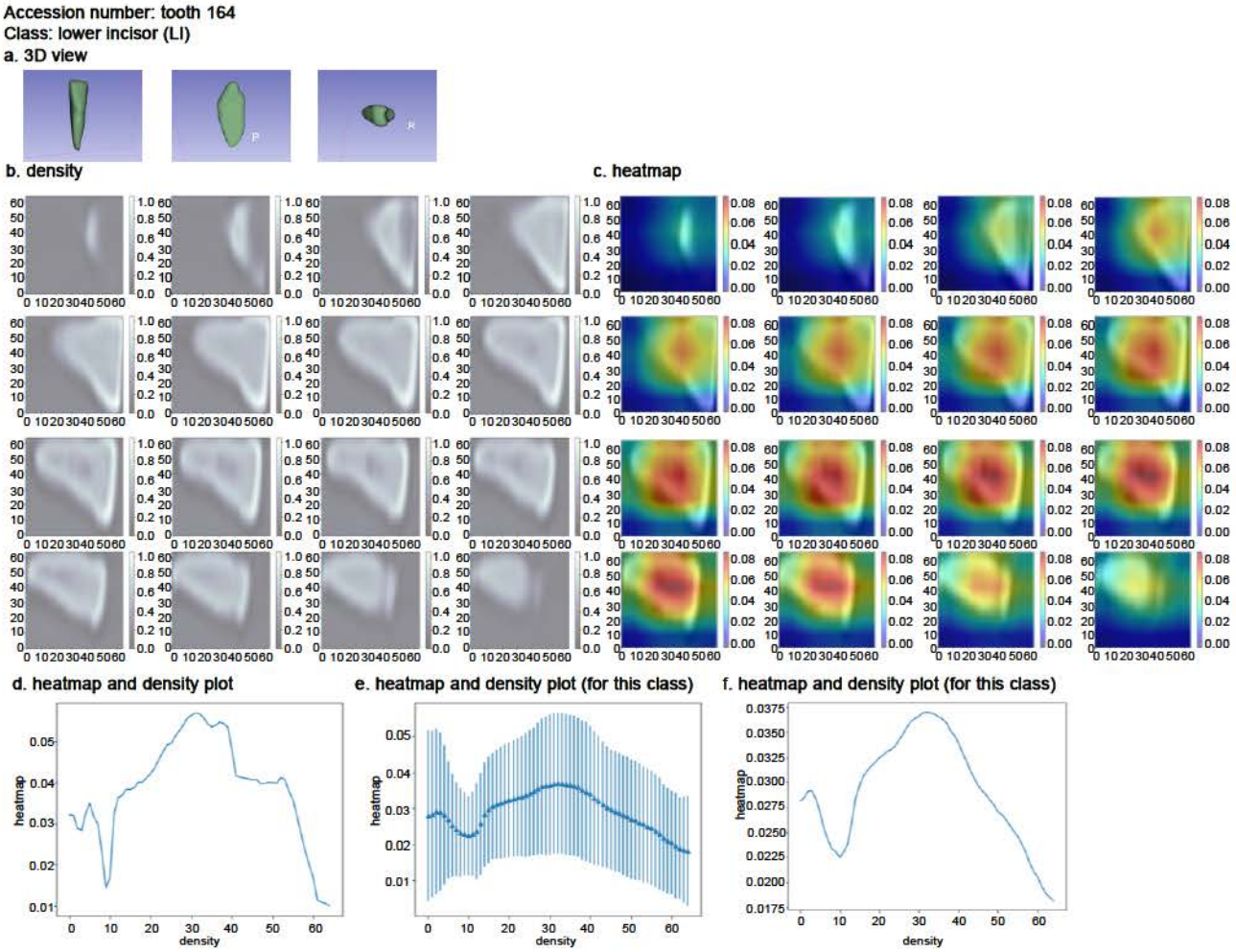
Supplementary File: Video demonstration of the interactive nature of the Machinometrics plot



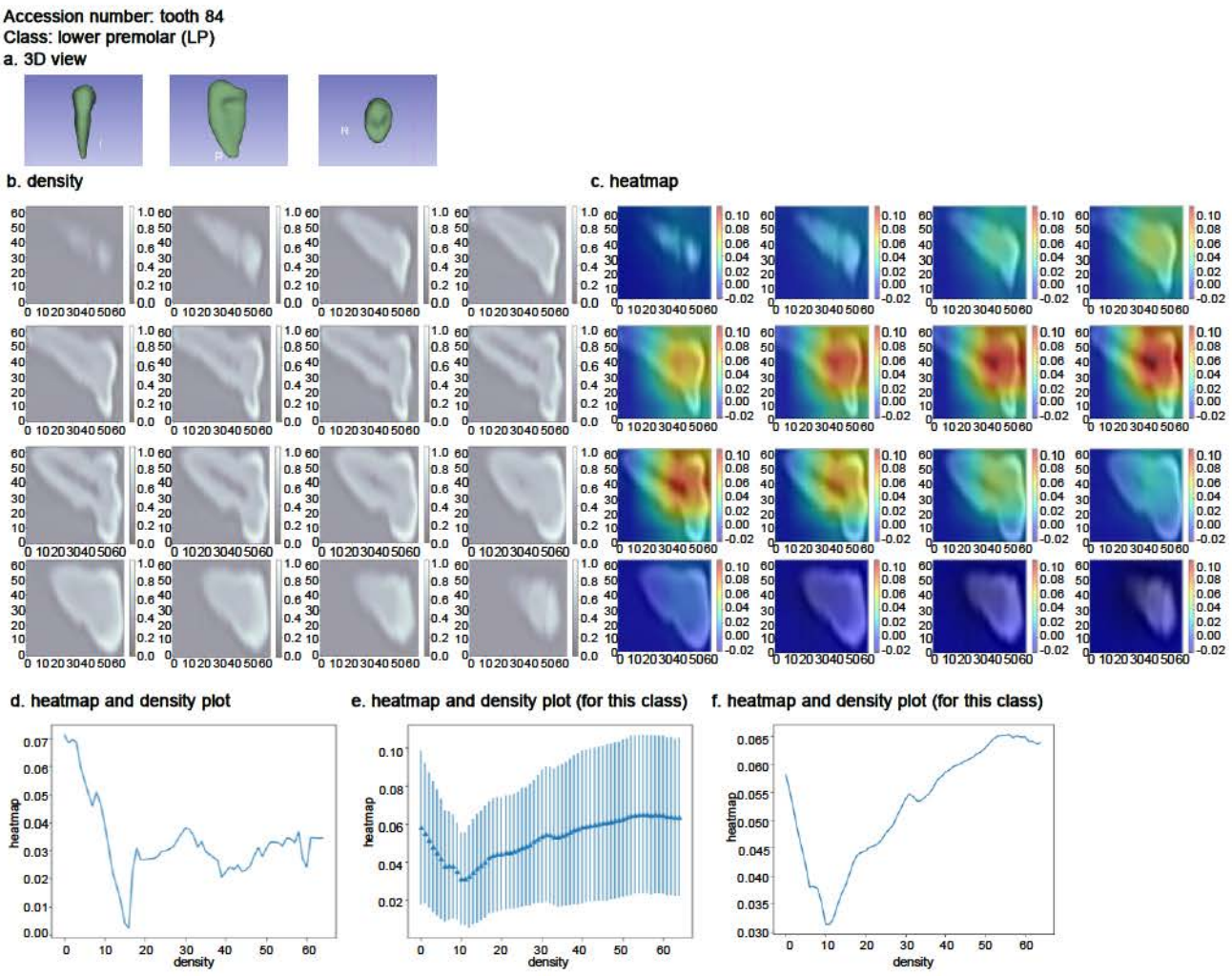
Extended Data File Figure 1: Representative CT scanned teeth shown before isolation from the background.



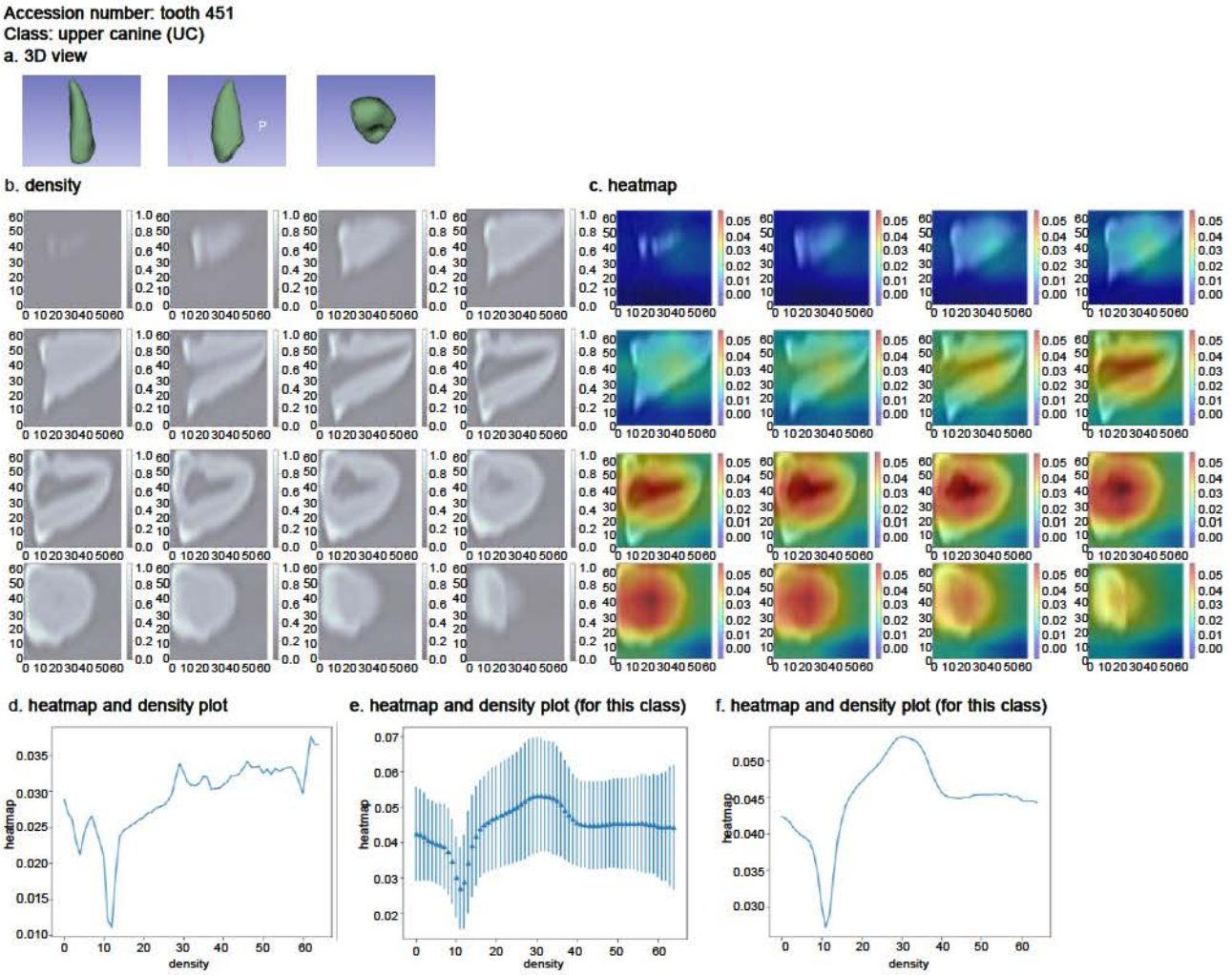
Extended Data File Figure 1.1: *Machinometrics* plots by integrating density for teeth class LC.



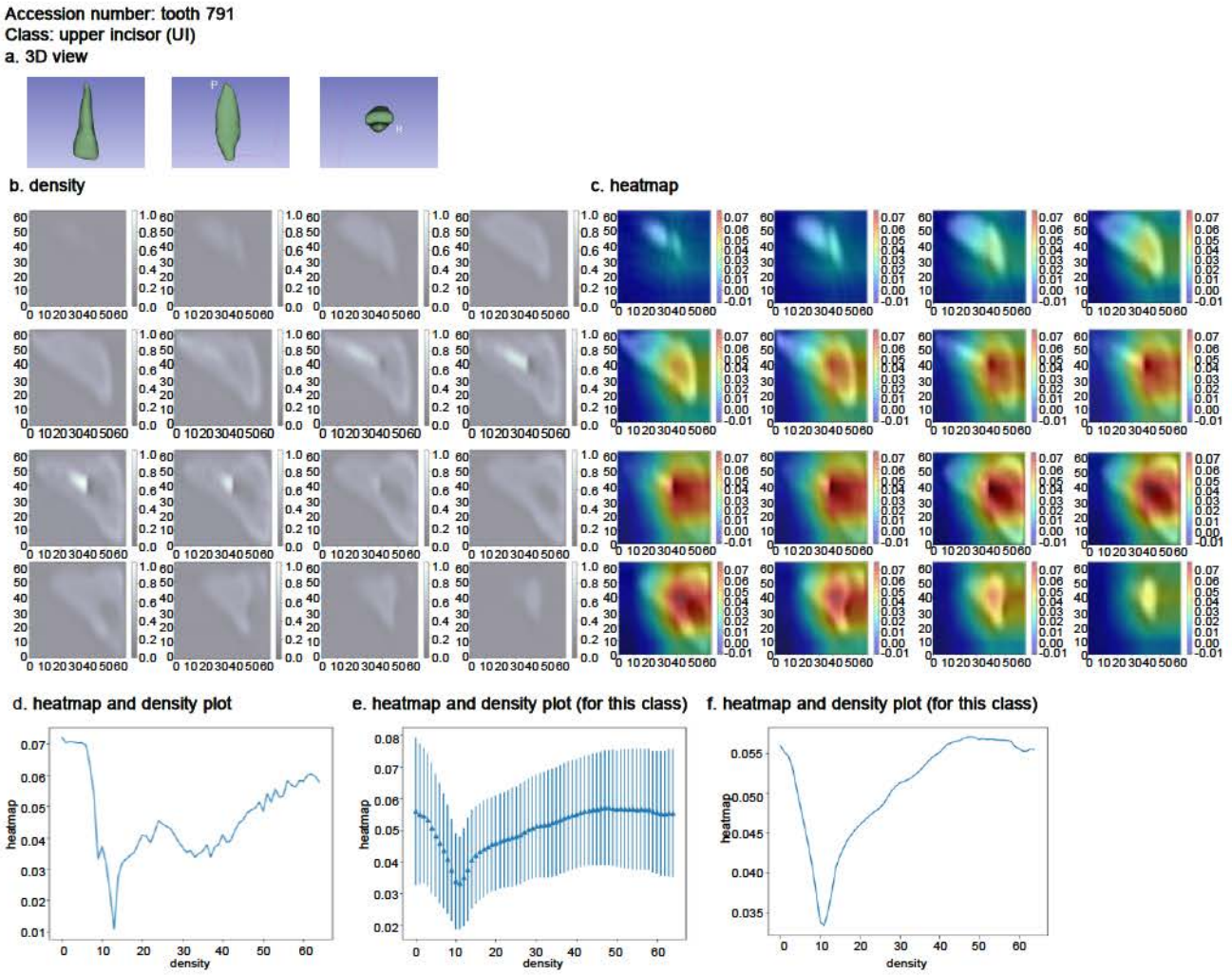
Extended Data File Figure 1.2: *Machinometrics* plots by integrating density for teeth class LI.



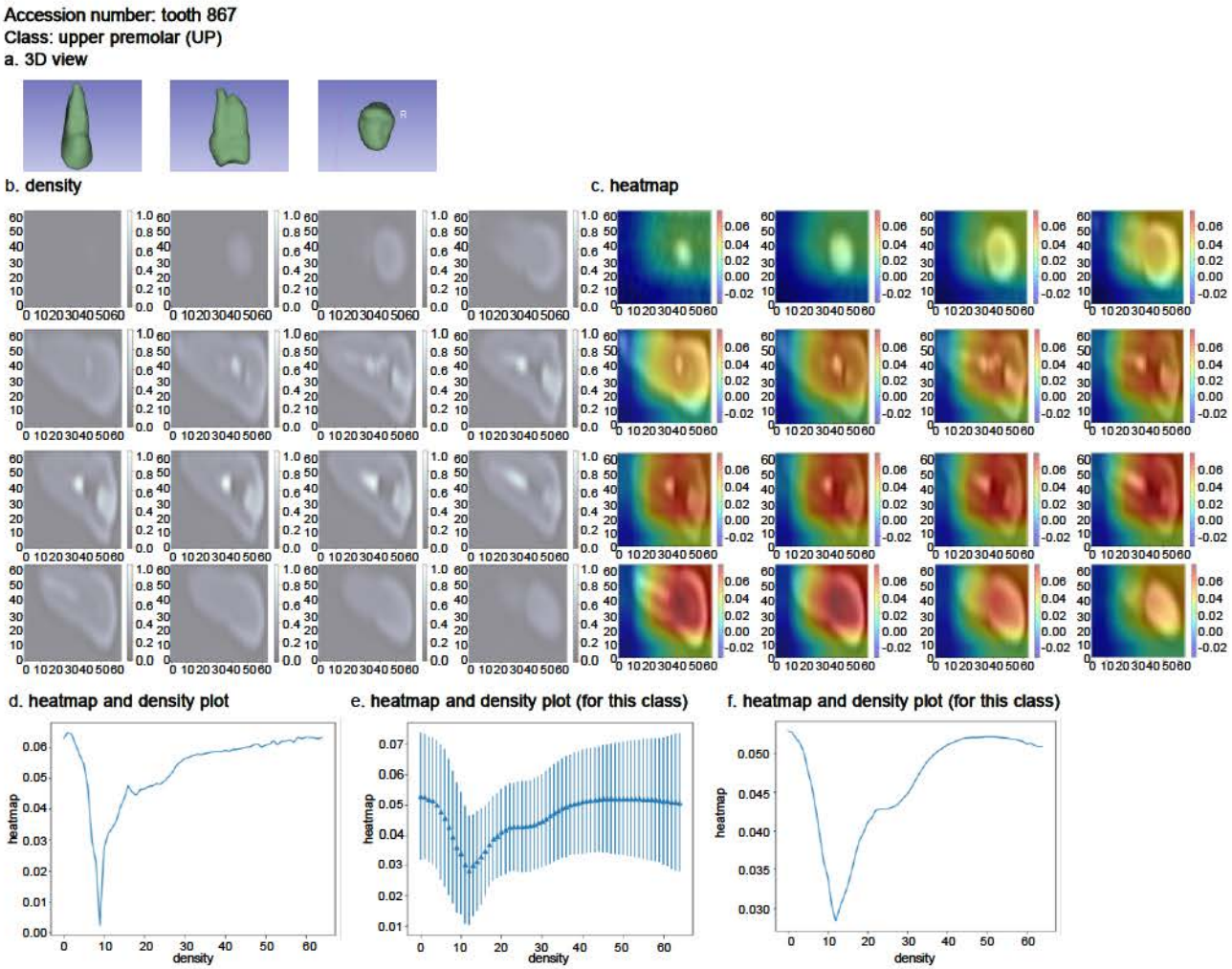
Extended Data File Figure 1.3: *Machinometrics* plots by integrating density for teeth class LP.



Extended Data File Figure 1.4: *Machinometrics* plots by integrating density for teeth class UC.

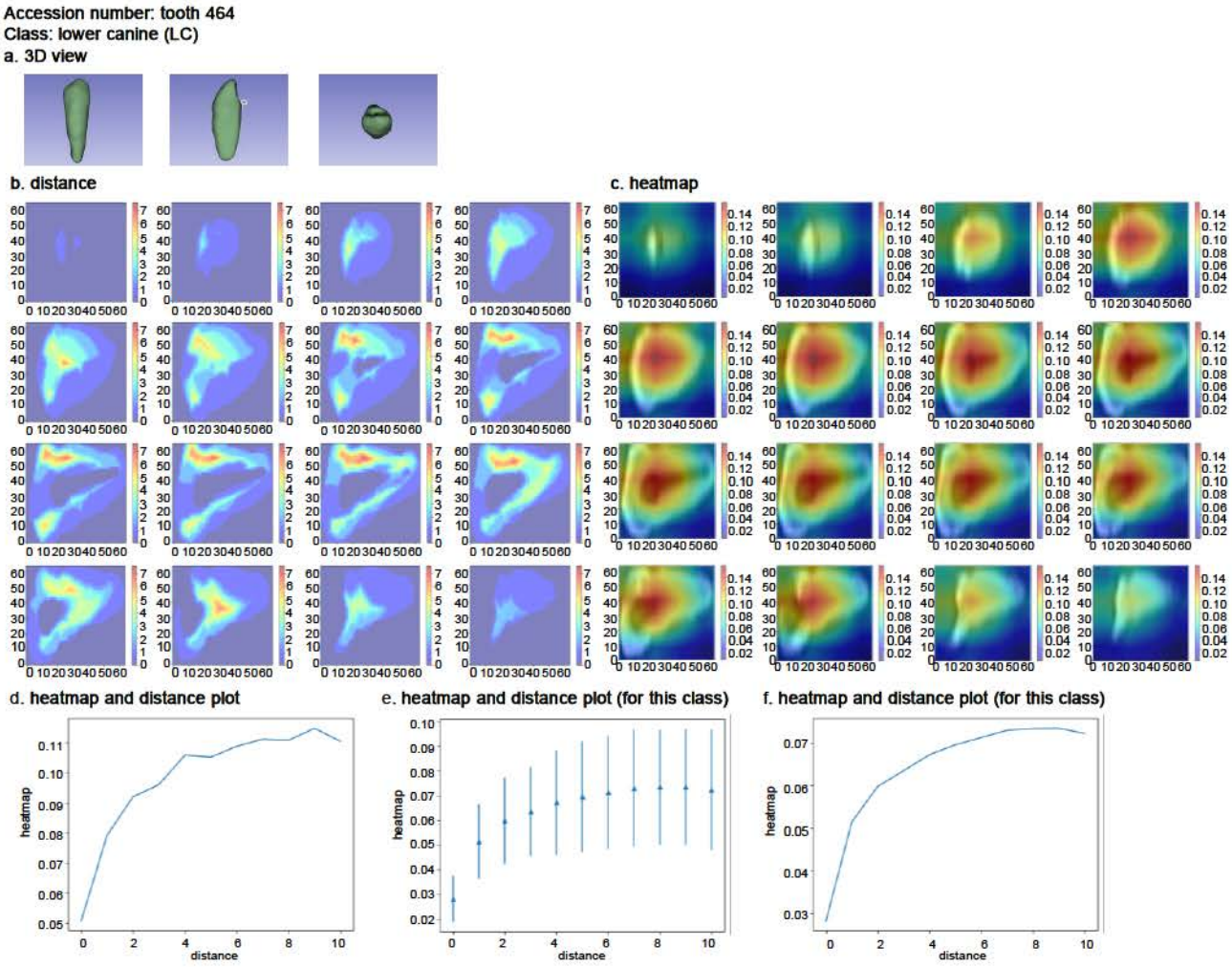


Extended Data File Figure 1.5: *Machinometrics* plots by integrating density for teeth class UI.



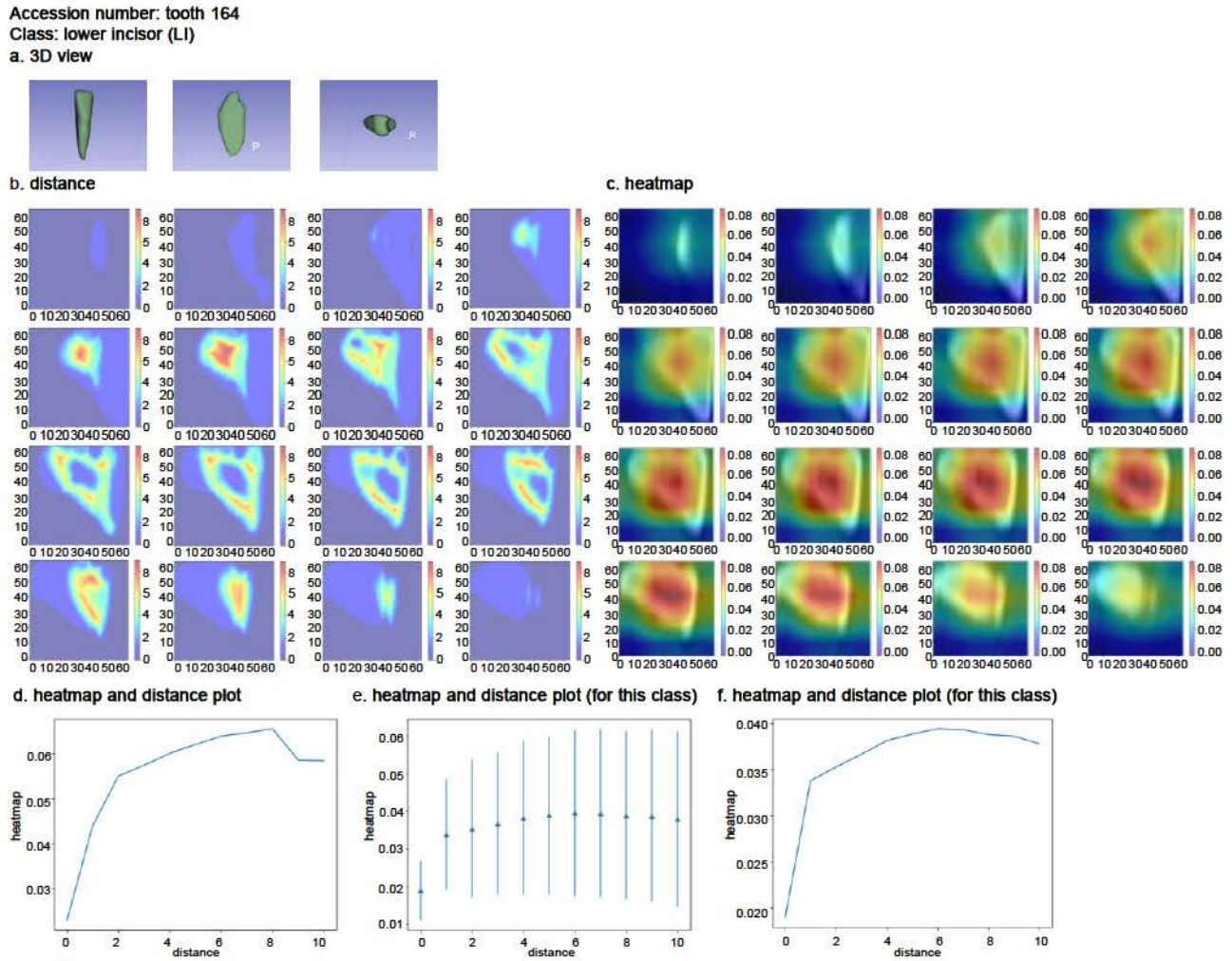
Extended Data File Figure 1.6: *Machinometrics* plots by integrating density for teeth class UP.

Extended Data File Figure 1.7



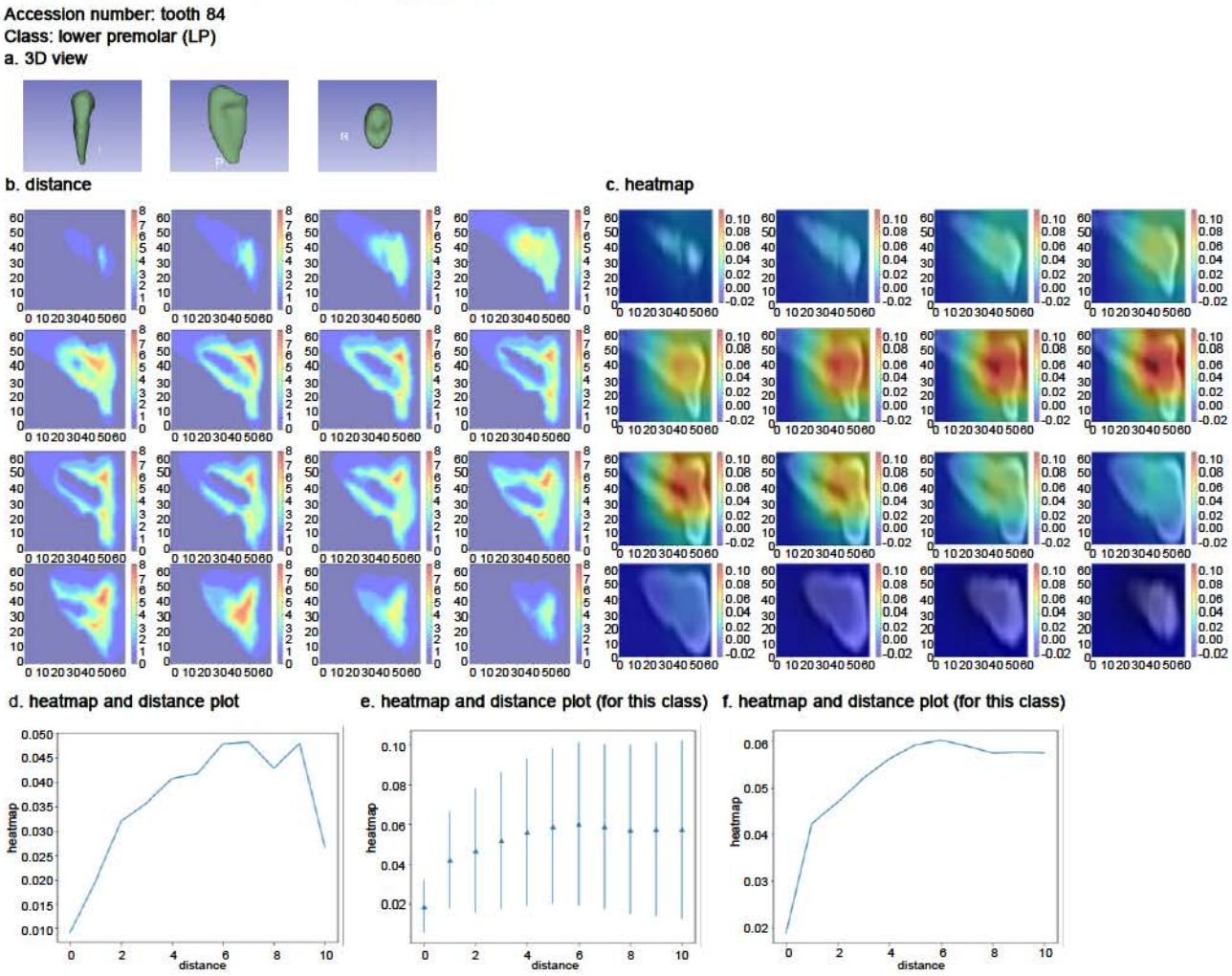
Extended Data File Figure 1.7: *Machinometrics* plots by integrating distance for teeth class LC.

Extended Data File Figure 1.8



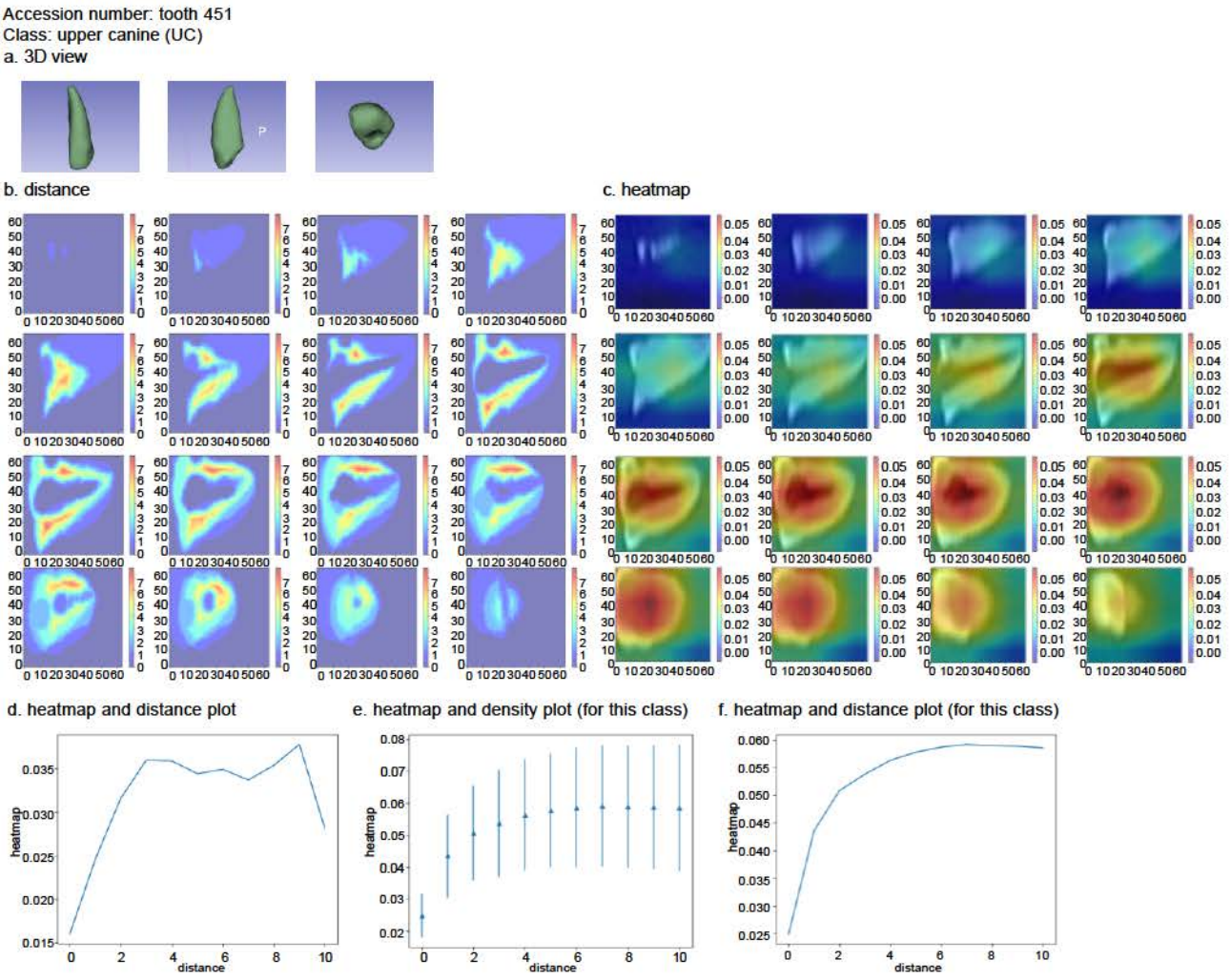
Extended Data File Figure 1.8: *Machinometrics* plots by integrating distance for teeth class LI .

Extended Data File Figure 1.9



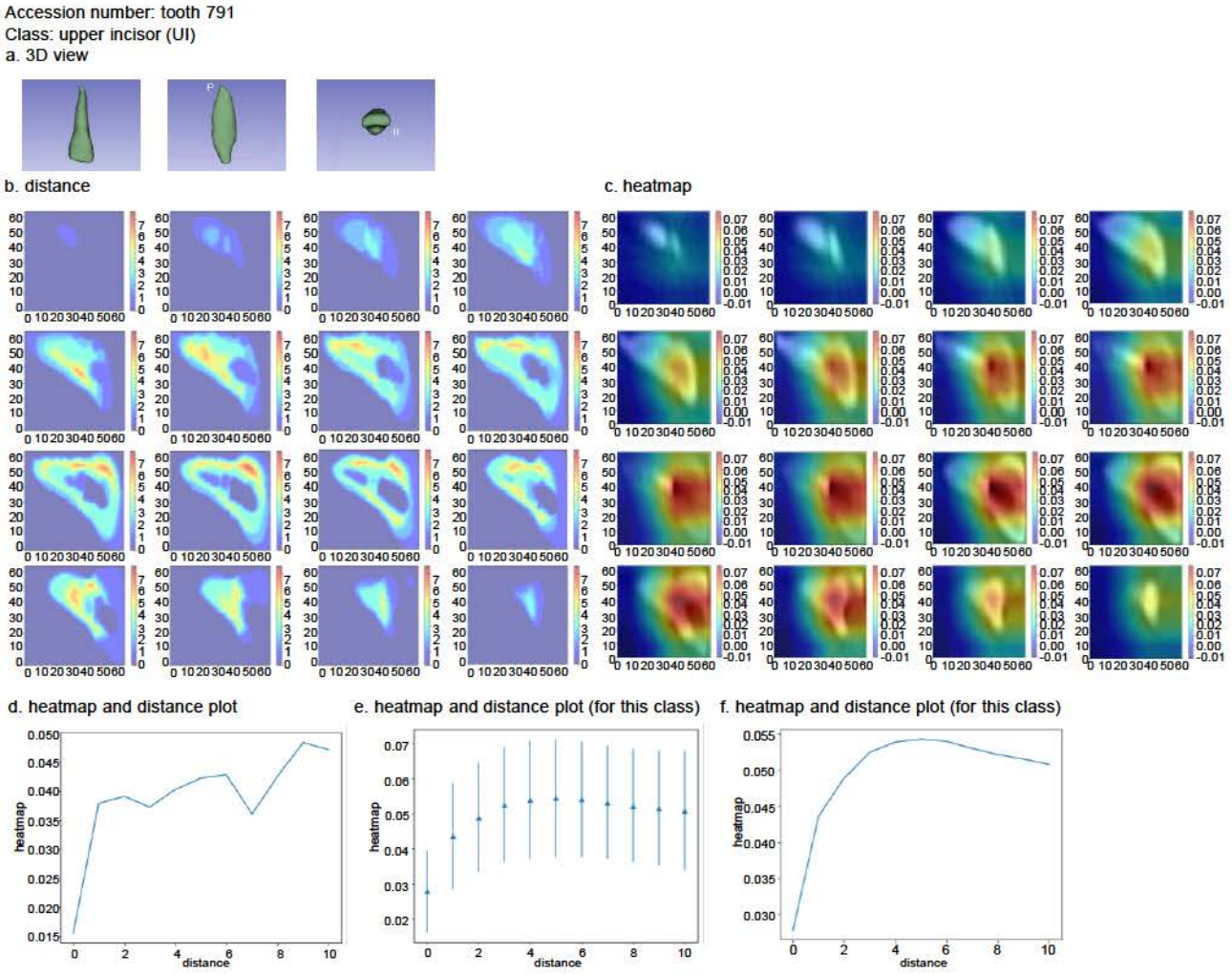
Extended Data File Figure 1.8: *Machinometrics* plots by integrating distance for teeth class LP .

Extended Data File Figure 1.10



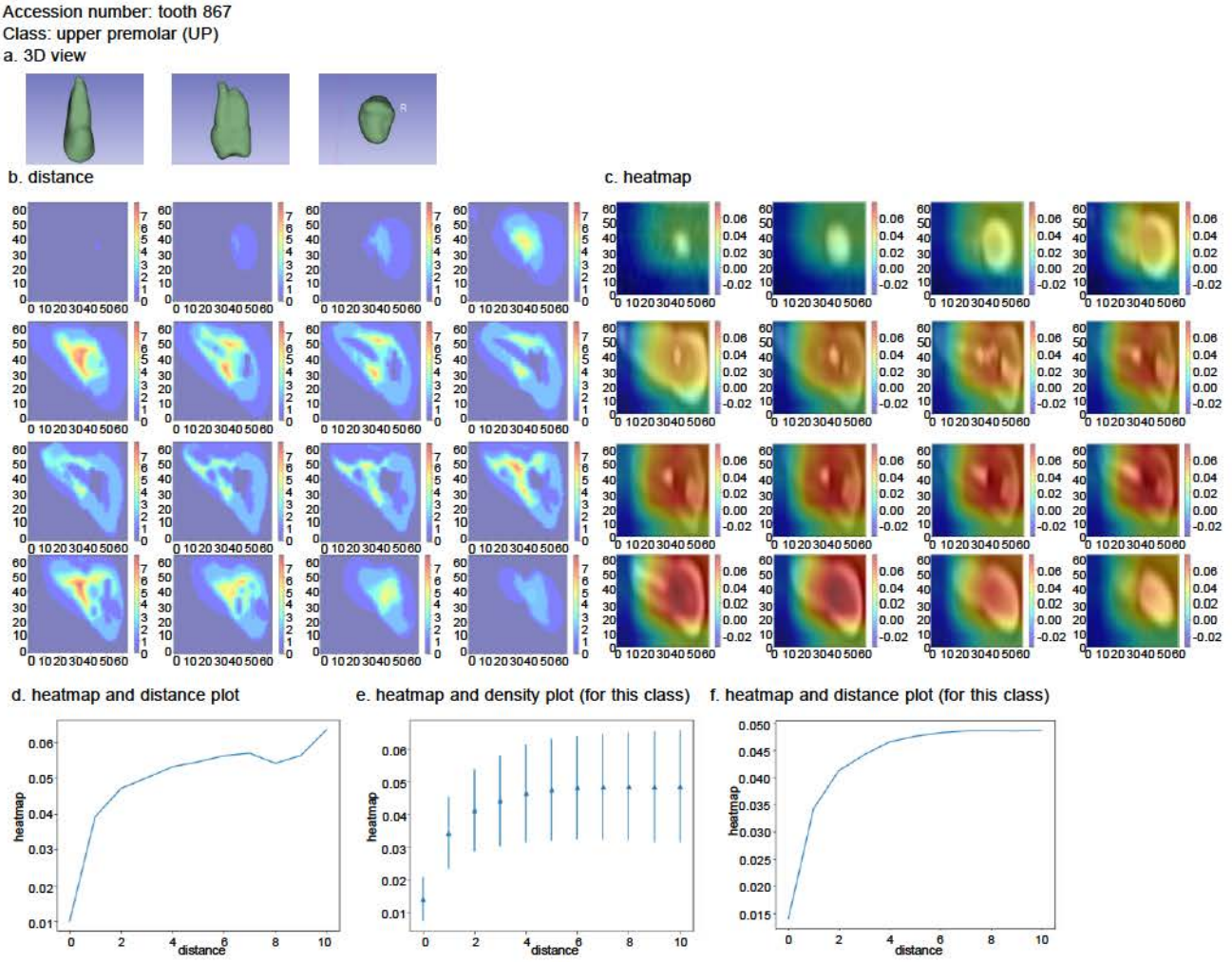
Extended Data File Figure 1.8: *Machinometrics* plots by integrating distance for teeth class UC .

Extended Data File Figure 1.11

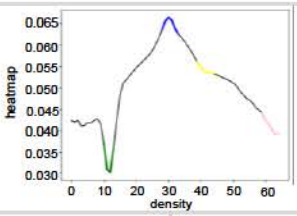
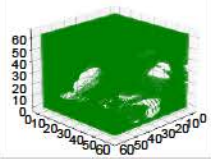
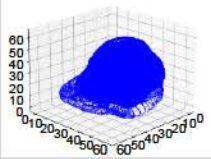
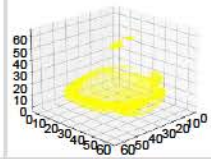
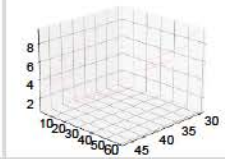
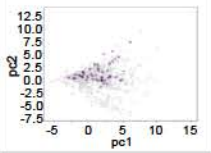
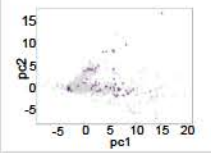
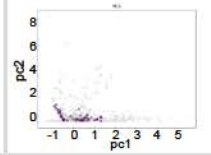
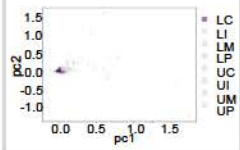


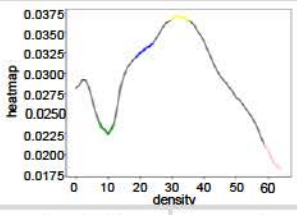
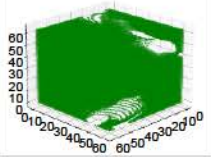
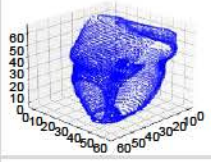
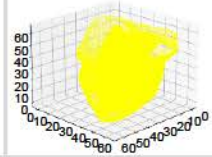
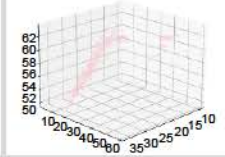
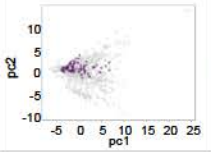
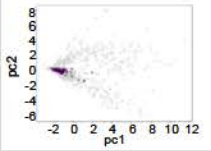
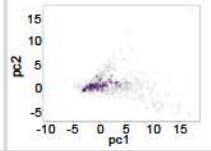
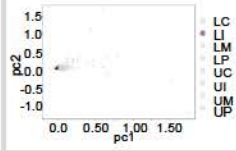
Extended Data File Figure 1.11: *Machinometrics* plots by integrating distance for teeth class UI.

Extended Data File Figure 1.12

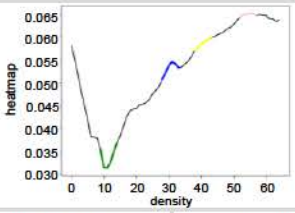
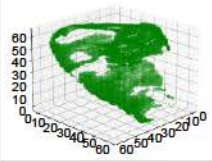
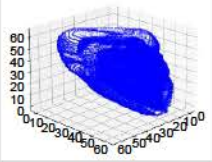
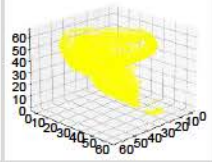
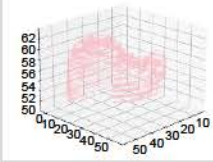
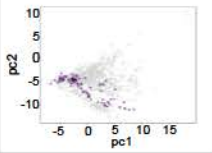
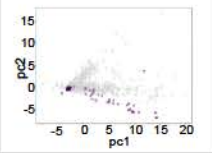
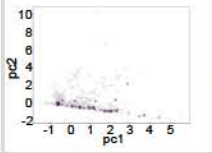
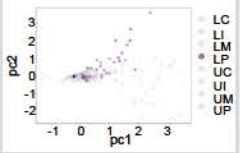


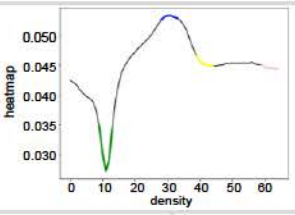
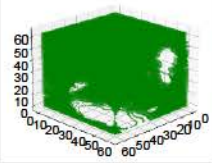
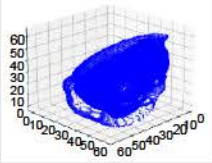
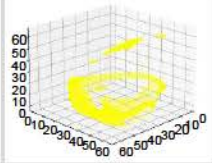
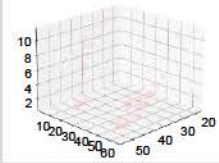
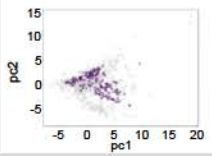
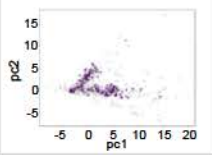
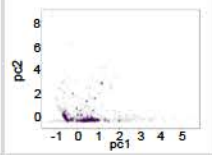
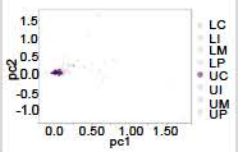
Extended Data File Figure 1.12: *Machinometrics* plots by integrating distance for teeth class UP.

Accession number: tooth 464, class: lower canine (LC)					
a1. heatmap and density plot					
b1. 3D visualisation		range from 10-13	range from 28-33	range from 39-44	range from 59-64
					
c1. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.14	0.24	0.05	0.03
	pc2	0.05	0.04	0.03	0.02
	Total	0.19	0.28	0.08	0.05

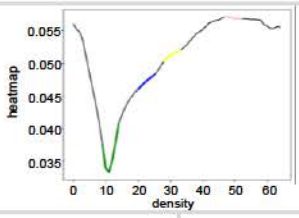
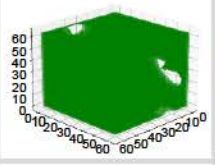
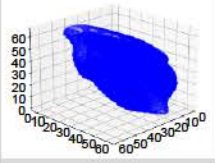
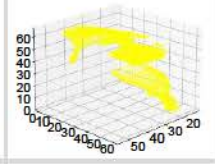
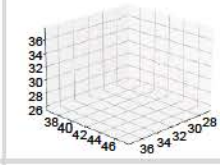
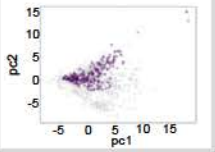
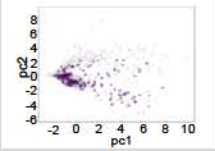
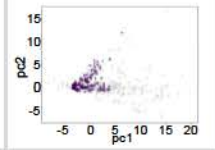
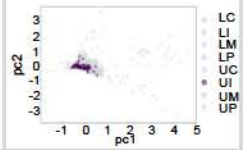
Accession number: tooth 161, class: lower incisor (LI)					
a2. heatmap and density plot					
b2. 3D visualisation		range from 7-12	range from 19-24	range from 30-35	range from 59-64
					
c2. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.18	0.11	0.22	0.03
	pc2	0.08	0.05	0.04	0.02
	Total	0.26	0.16	0.26	0.05

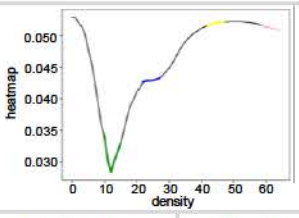
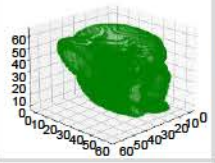
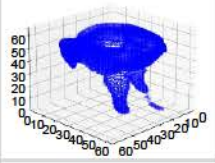
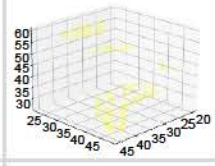
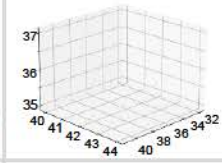
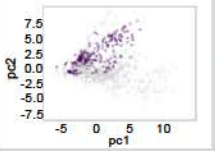
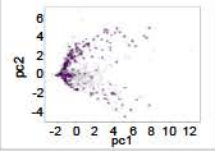
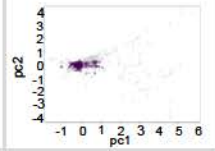
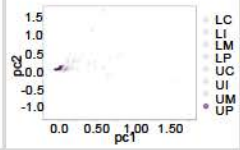
Extended Data File figures 2.1: 3D visualisations for density ranges for teeth classes LC, and LI.

Accession number: tooth 84, class: lower premolar (LP)					
a1. heatmap and density plot					
b1. 3D visualisation		range from 9-14	range from 28-33	range from 38-43	range from 52-57
					
c1. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.16	0.22	0.05	0.08
	pc2	0.07	0.04	0.04	0.02
	Total	0.23	0.26	0.09	0.10

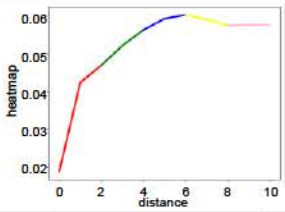
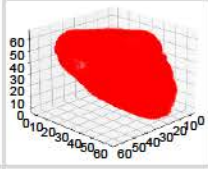
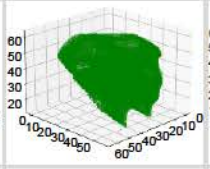
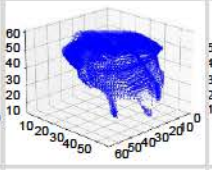
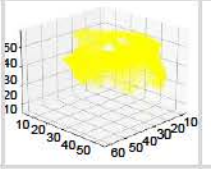
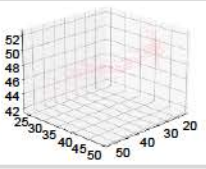
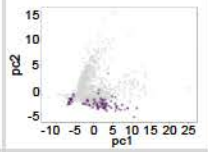
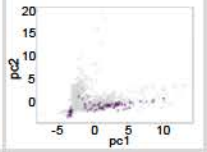
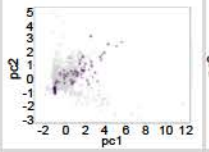
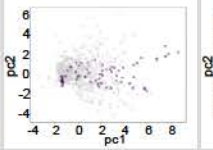
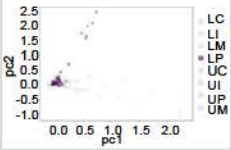
Accession number: tooth 451, class: upper canine (UC)					
a2. heatmap and density plot					
b2. 3D visualisation		range from 9-13	range from 28-33	range from 39-44	range from 59-64
					
c2. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.16	0.24	0.05	0.03
	pc2	0.07	0.04	0.03	0.02
	Total	0.23	0.28	0.08	0.05

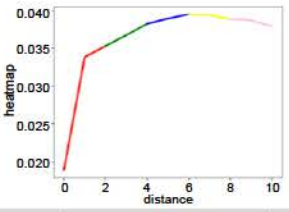
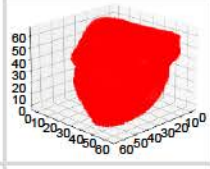
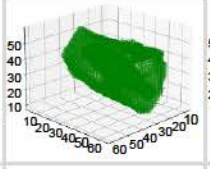
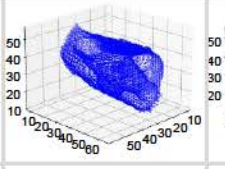
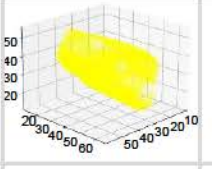
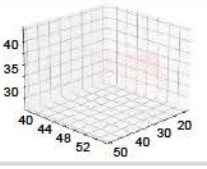
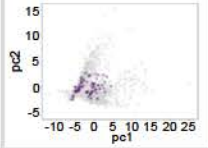
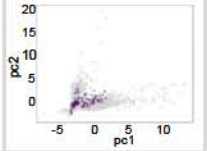
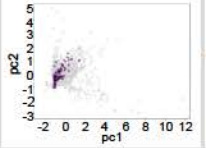
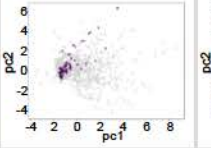
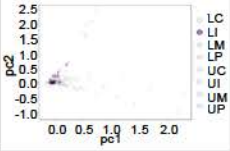
Extended Data File figure 2.2: 3D visualisations for density ranges for teeth classes LP and UC.

Accession number: tooth 791, class: upper incisor (UI)					
a1. heatmap and density plot					
b1. 3D visualisation		range from 9-14	range from 20-25	range from 28-33	range from 47-52
					
c1. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.16	0.10	0.24	0.08
	pc2	0.07	0.05	0.04	0.03
	Total	0.23	0.15	0.28	0.11

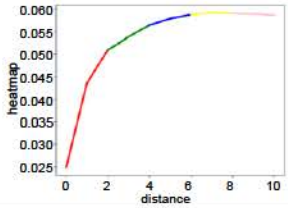
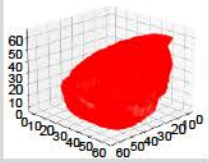
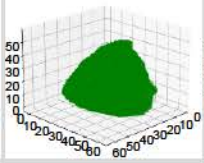
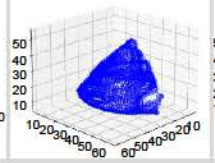
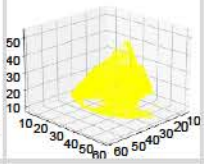
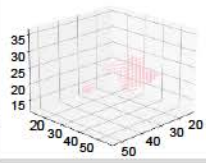
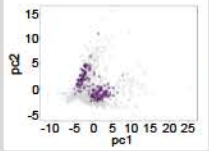
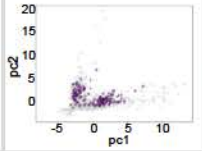
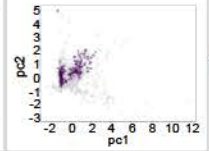
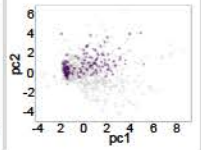
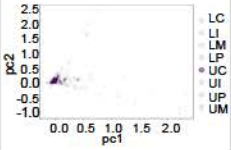
Accession number: tooth 867, class: upper premolar (UP)					
a2. heatmap and density plot					
b2. 3D visualisation		range from 0-5	range from 22-27	range from 42-47	range from 59-64
					
c2. principal component analysis (PCA)					
Variant explained by PCA	pc1	0.14	0.07	0.07	0.03
	pc2	0.06	0.04	0.02	0.02
	Total	0.20	0.11	0.09	0.05

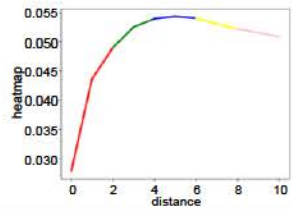
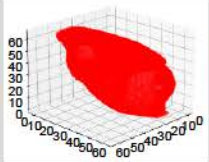
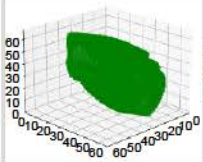
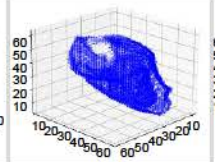
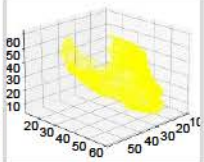
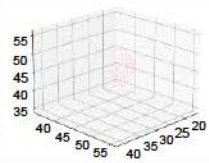
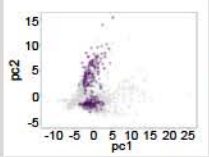
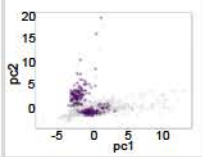
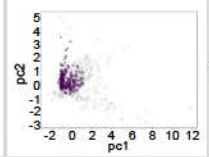
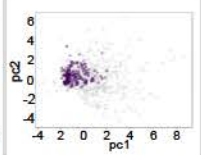
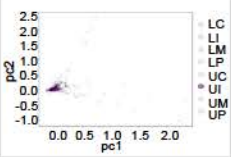
Extended Data File Figure 2.3: 3D visualisations for density ranges for teeth classes UI, and UP.

Accession number: tooth 84, class: lower premolar (LP)						
a1. heatmap and distance plot						
b1. 3D visualisation	range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10	
						
c1. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

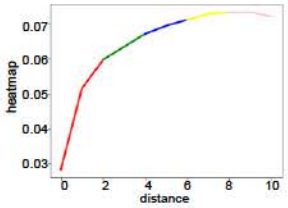
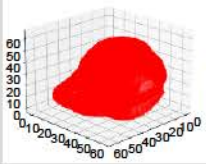
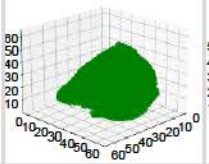
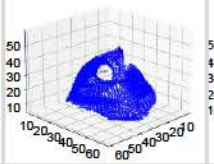
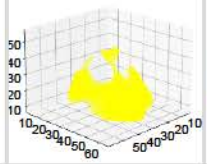
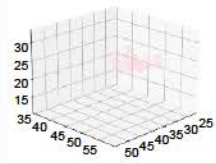
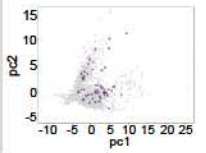
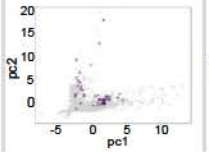
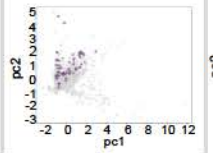
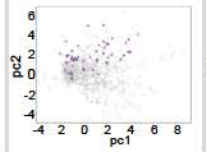
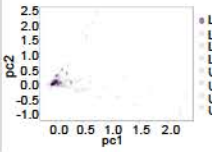
Accession number: tooth 164, class: lower incisor (LI)						
a2. heatmap and distance plot						
b2. 3D visualisation	range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10	
						
c2. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

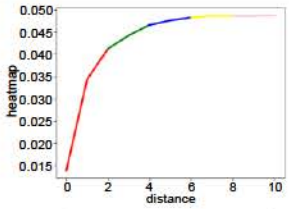
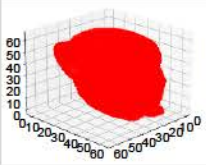
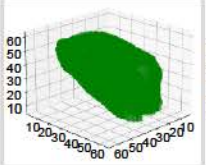
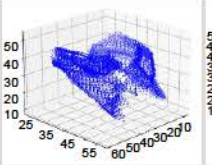
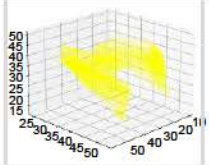
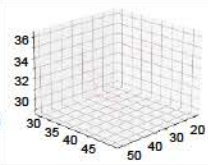
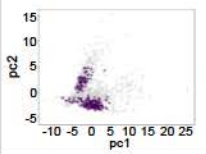
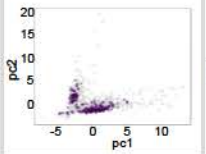
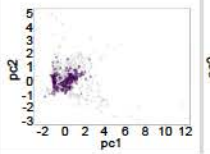
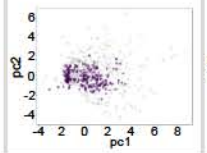
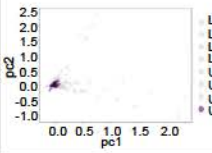
Extended Data File Figure 2.4: 3D visualisations for distance ranges for teeth classes LP, and LI.

Accession number: tooth 451, class: upper canine (UC)						
a1. heatmap and distance plot						
b1. 3D visualisation		range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10
						
c1. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

Accession number: tooth 791, class: upper incisor (UI)						
a2. heatmap and distance plot						
b2. 3D visualisation		range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10
						
c2. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

Extended Data File Figure 2.5: 3D visualisations for distance ranges for teeth classes UC, and UI.

Accession number: tooth 464, class: lower canine (LC)						
a1. heatmap and distance plot						
b1. 3D visualisation	range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10	
						
c1. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

Accession number: tooth 867, class: upper premolar (UP)						
a2. heatmap and distance plot						
b2. 3D visualisation	range from 0-2	range from 2-4	range from 4-6	range from 6-8	range from 8-10	
						
c2. principal component analysis (PCA)						
Variant explained by PCA	pc1	0.11	0.11	0.06	0.11	0.03
	pc2	0.06	0.05	0.02	0.05	0.03
	Total	0.17	0.16	0.08	0.16	0.06

Extended Data File Figure 2.6: 3D visualisations for distance ranges for teeth classes LC, and UP.

Chapter 8:

General Discussion and Conclusion

This thesis explored how machine learning algorithms can serve as cognitive aids to reduce interpretive errors in dental radiology in the context of radiographic caries diagnosis. The findings of this work have implications for improving the accuracy of radiologic diagnosis in dentistry.

8.1 Analysis of factors contributing to interpretive errors

This research revealed that errors of omission were perceived as the most prevalent interpretive errors in dental radiology. Errors of omission occur due to the clinician failing to 'see' or recognise the lesion on the radiograph or incorrectly identifying the significance of the lesion.[1,2] This finding in our survey contrasted what is known in medical radiology, where delayed diagnosis and misdiagnosis are seen as the most commonly occurring types of error.[3]

The study identified several factors contributing to these errors, including clinicians' experience, technical skills, time pressure, and cognitive load. Data suggests that interpretive errors resulted primarily from increased cognitive load due to factors such as excessive workload, lack of clinical experience, and poor image quality. Further, image interpretation related factors such as jumping to conclusions regarding diagnosis and misdiagnosis due to a lesion mimicking another lesion were also seen as likely causes of these errors. These findings were supported by the systematic review of literature and the survey of dental practitioners in NSW. These errors can have serious consequences, such as delayed diagnosis, inappropriate treatment, and increased patient morbidity.

While our research found that dental practitioners preferred traditional methods such as continuing dental education and peer consultation as methods to minimise errors, there is growing evidence that machine learning can significantly enhance diagnostic accuracy and improve dental treatment outcomes.[4–6] Addressing the concerns regarding data privacy, output accuracy, and transparency of AI decisions is essential for wider acceptance of AI technologies.[7] Our study found that assistance from ML algorithms not only increased the accuracy of diagnosis of dental caries on radiographs but also boosted student clinician confidence in their diagnosis. This increase in confidence could indirectly lead to reduced cognitive load, as reduced uncertainty in diagnosis could lower the mental effort required in diagnosis. In dentistry, clinicians often face high cognitive demands due to the complexity of images, time pressures, and the need to integrate vast amounts of clinical information.[8] Excessive cognitive load can impair judgement, leading to errors such as missing a lesion (error of omission) or misinterpreting radiographic findings.[9]

8.2 Role of cognitive aids in improving diagnostic accuracy

Understanding the complexities of how humans make decisions, especially in the context of clinical decision-making, is challenging. Clinicians develop tacit knowledge while analysing images, which includes having a consistent gaze trajectory, an intuitive understanding of the measurements, location, and size of findings on the images and other habits acquired over time to inform them of the probability of pathology on an image.[10] Knowledge of the clinical case, the patient's chief complaint, and the patient's history also play a role in the decision-making process.[11] Tacit knowledge is subjective, experience-based knowledge or skills or competencies that are usually context-specific and cannot be articulated. Explicit knowledge, on the other hand, is objective, rational knowledge that is expressed as words, sentences, numbers, etc. It includes theoretical approaches, problem-solving and manuals.[12] There are significant benefits to making tacit knowledge explicit, as it can be advantageous at both individual and organisational levels. It can also facilitate effective training and education programs by formalising the knowledge and experience of experienced clinicians. Making tacit knowledge explicit can also facilitate the development of ML clinical decision support systems that align with clinical best practices, leading to enhanced patient care.[13] When experienced clinicians annotate radiographic images, they externalise their intuitive expertise, transforming tacit knowledge into explicit data that can be used to train ML algorithms. Further, when clinicians interact with ML algorithms, the algorithm can highlight features, patterns or anomalies on the radiographs. This can prompt the clinicians to reflect on their own decision-making process and effectively externalise their tacit knowledge. [14] This can have several implications, such as improved diagnostic accuracy, helping novice clinicians perform better, and providing a continuous feedback loop for clinicians.

Clinical situations are unpredictable and are characterised by many known and unknown variables, and clinicians may use both System 1 and 2 simultaneously to arrive at a diagnosis.[15] System 1, or fast thinking, is also called heuristics and is rapid reflexive decision-making. Research shows that heuristics develop with clinical experience, and experienced clinicians may reach a diagnosis intuitively without deliberate thought (clinical gestalt)[15]. While System 1 thinking is efficient, it is prone to errors due to cognitive biases leading to errors. System 2 thinking, also called slow thinking, involves slow, deliberate decision-making. Inexperienced clinicians may use system 2 thinking in diagnosis.[16,17] While experienced clinicians often rely on intuitive judgements (system 1) to form initial hypotheses, challenging cases may require deliberate analysis (system 2) to validate their diagnosis.[17] ML algorithms can influence clinical decision-making by encouraging the shift from system 1 to system 2 thinking when needed. When ML algorithms highlight areas of concern on radiographs, they prompt clinicians to engage in more deliberate analysis, reducing the reliance on potentially error prone intuitive judgments. This shift can enhance diagnostic accuracy by mitigating the impact of cognitive load on the interpretation of radiographs. For example, ML can detect subtle radiolucencies indicative of early dental caries that a clinician might miss due to cognitive overload or perceptual limitations. In addition to supporting clinicians with clinical decision-making, ML algorithms have the potential to minimise interpretive errors. ML algorithms are trained on large datasets and can detect abnormalities with high precision, often detecting subtle findings that can be missed by clinicians.[18]

This research has shown that dental interpretive errors are influenced by various clinician-related and image interpretation related factors. In this context, ML algorithms can support clinicians by providing a second opinion, acting as a safety net, and reducing missed and misdiagnoses on dental radiographs. Additionally, research has shown that inter- and intra-observer variability can affect the diagnosis of common dental conditions.[19–21] ML can mitigate these variations, standardising the radiographic interpretive process and providing a more consistent and accurate diagnosis. Figure 3 depicts a diagrammatic representation of the collaborative role of ML algorithms in dental diagnosis.

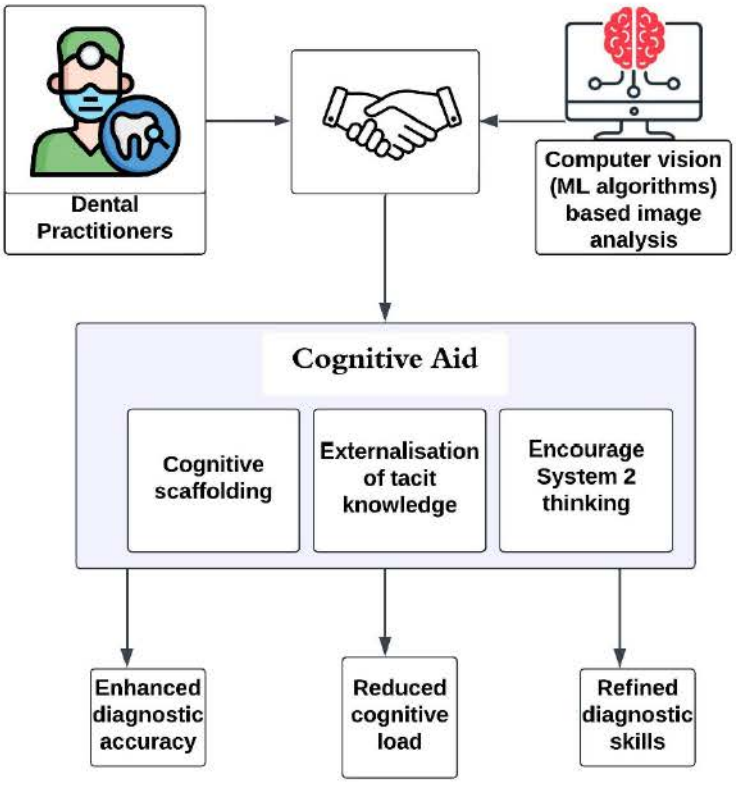


Figure 3: Diagrammatic representation of the collaborative role of ML algorithms

8.3 AI in dentistry: opportunities and challenges

ML algorithms could reduce cognitive load by providing objective analyses of radiographic images, free from human cognitive biases, facilitating this through mechanisms such as cognitive scaffolding and encouraging a shift to System 2 thinking.[22,23] ML algorithms provide cognitive scaffolding by reducing cognitive load, enhancing the clinician's ability to recognise pathologies on radiographs by consistently highlighting areas of concern on radiographs. Cognitive scaffolding is a method of providing structured support by which the complexity of a task is reduced. Reducing the complexity allows for cognitive resources of working memory to be allocated more efficiently, improving accuracy.[23] ML algorithms may also encourage a shift from a more intuitive System 1 to an analytic System 2 thinking. Thus, they can help

clinicians externalise tacit knowledge by providing reminders or acting as an assistive tool. Because AI systems are continuously learning from expert annotations and by analysing large datasets, AI can offer recommendations that reflect the tacit knowledge and patterns recognised by experienced clinicians.[24] When AI systems are used by novice clinicians, they can ensure consistency and accuracy in diagnosis. This collaborative interaction between human clinicians and ML systems exemplifies effective cognitive scaffolding without replacing the clinician's essential judgment and expertise. [25]

Integrating artificial intelligence technologies into dental practice represents a significant advancement with the potential to enhance diagnostic accuracy, treatment planning and overall patient care.[26] However, the successful appropriation of AI in dentistry depends on a mutual influence between technology and clinicians. This reciprocal relationship emphasises that technology must not only be adopted by dental professionals but also adapted so that it fits seamlessly within their established clinical workflows. For AI technologies to be effectively and widely adopted, they must align with the professional expectations and practices of dentists. AI tools must support clinical guidelines, ethical considerations and patient-centred care philosophies. To achieve this, the design and development of AI technologies must be iterative, transparent, and participative, and they must act as a collaborative tool rather than replace human judgement. This research has suggested a novel method to accomplish the transparent, explainable design of machine learning algorithms.

This collaborative relationship can be characterised by concepts of openness, multiplicity and auxiliarity.[27] AI systems should be able to integrate within the existing dental practice management systems and be adaptable to various clinical settings. Recognising that there are multiple ways to approach diagnosis and treatment planning, AI tools should accommodate diverse clinical perspectives and methodologies. Multiplicity ensures that AI tools support a range of diagnostic strategies rather than enforcing a single approach. Auxiliarity emphasises that AI is a smart assistant and supports clinicians in decision-making.[28] This flexibility ensures that ML algorithms can complement the clinician's decision-making process rather than undermine it. By incorporating these principles, AI technologies can be effectively appropriated within dental practice. This leads to a relationship where technology evolves based on clinical needs, and clinicians enhance their practice through capabilities provided by AI. Ultimately, the success of AI integration in dentistry depends on its ability to respect the professional autonomy of dentists while providing tangible benefits to improve patient outcomes.

Conclusions:

8.4 Summary of Key Findings

This research highlights the impact of machine learning algorithms in improving diagnostic accuracy during the interpretation of dental radiographs. By alleviating the cognitive load on clinicians, ML algorithms allow for more focused and effective decision-making. They can provide objective, consistent analyses that can mitigate human cognitive biases and be viewed as collaborative tools rather than replace human judgements. To successfully implement AI

systems into clinical workflows, AI systems must become more explainable and transparent, providing insights that align with clinical reasoning. Clinicians, on the other hand, need to develop a better understanding of how AI systems work and their limitations. This can be achieved by a collaborative approach, where AI and the clinician work together.

This research is highly relevant to the wider dental profession, including dental students, dentists, educators and policymakers. For students, it emphasises the importance of integrating advanced technologies into their learning to enhance diagnostic accuracy and prepare for future practice. For the dental profession, it highlights the potential benefits of ML algorithms in improving patient outcomes, streamlining clinical workflows and fostering innovation in clinical practice.

8.5 Recommendations for Future Research

The findings of this research open several avenues for further investigation and development. Future studies will focus on investigating how ML algorithms can influence different aspects of clinical decision-making, not just accuracy but also speed, consistency and confidence with which decisions are made in clinical situations. Further research could also investigate how ML algorithms can influence the development of diagnostic skills in dental students and novice clinicians. Do ML algorithms help novices better identify and learn from errors? Such studies could also investigate how to best integrate ML algorithms into educational curricula in dentistry.

Another key area of future research is to explore different human-AI collaboration models and investigate how ML algorithms can be incorporated into clinical workflows to minimise interpretive errors.

8.6 References:

1. Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. World J Radiol [Internet]. 2010 Oct;2(10):377–83. Available from: http://www.wjgnet.com/1949-8470/full/v2/i10/377.htm
2. Adamo SH, Gereke BJ, Shomstein S, Schmidt J. From "satisfaction of search" to "subsequent search misses": a review of multiple-target search errors across radiology and cognitive science. Cogn Res Princ Implic [Internet]. 2021;6(1):59. Available from: https://doi.org/10.1186/s41235-021-00318-w
3. Kim, YW ML. Fool me Twice: Delayed Diagnosis in radiology. Am J Roentgenol. 2014;202:465–70.
4. Khanagar SB, Al-ehaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, et al. Developments, application, and performance of artificial intelligence in dentistry – A systematic review. J Dent Sci [Internet]. 2021;16(1):508–22. Available from: https://www.sciencedirect.com/science/article/pii/S1991790220301434

5. Al-Namankany A. Influence of Artificial Intelligence-Driven Diagnostic Tools on Treatment Decision-Making in Early Childhood Caries: A Systematic Review of Accuracy and Clinical Outcomes. Vol. 11, Dentistry Journal. 2023.

6. Hung KF, Yeung AWK, Bornstein MM, Schwendicke F. Personalised dental medicine, artificial intelligence, and their relevance for dentomaxillofacial imaging. Dentomaxillofac Radiol. 2023 Jan;52(1):20220335.

7. Williamson SM, Prybutok V. Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. Vol. 14, Applied Sciences. 2024.

8. Feller L, Lemmer J, Nemutandani MS, Ballyram R, Khammissa RAG. Judgment and decision-making in clinical dentistry. J Int Med Res [Internet]. 2020 Nov 1;48(11):0300060520972877. Available from: https://doi.org/10.1177/0300060520972877

9. Westbrook JI, Raban MZ, Walter SR, Douglas H. Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: a prospective, direct observation study. BMJ Qual Saf. 2018 Aug;27(8):655–63.

10. Anderson B, Shyu CR. A preliminary study to understand tacit knowledge and visual routines of medical experts through gaze tracking. AMIA . Annu Symp proceedings AMIA Symp. 2010 Nov;2010:21–5.

11. Hajjaj FM, Salek MS, Basra MKA, Finlay AY. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med. 2010 May;103(5):178–87.

12. Nonaka I. A Dynamic Theory of Organizational Knowledge Creation. Organ Sci [Internet]. 1994 Sep 29;5(1):14–37. Available from: http://www.jstor.org/stable/2635068

13. van Baalen S, Boon M, Verhoef P. From clinical decision support to clinical reasoning support systems. J Eval Clin Pract. 2021 Jun;27(3):520–8.

14. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020 Mar;46(3):205–11.

15. Croskerry P. A Universal Model of Diagnostic Reasoning. Acad Med [Internet]. 2009;84(8). Available from: https://journals.lww.com/academicmedicine/Fulltext/2009/08000/A_Universal_Model_of_Diagnostic_Reasoning.14.aspx

16. Tay SW, Ryan P, Ryan CA. Systems 1 and 2 thinking processes and cognitive reflection testing in medical students. Can Med Educ J. 2016 Oct;7(2):e97–103.

17. Benishek LE, Weaver SJ, Newman-Toker DE. The Cognitive Psychology of Diagnostic Errors. Sci Am Neurol. 2015;(10).

18. Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual and Interpretive Error in Diagnostic Radiology—Causes and Potential Solutions. Acad

Radiol [Internet]. 2019;26(6):833–45. Available from: https://www.sciencedirect.com/science/article/pii/S107663321830521X

19. Meusburger T, Wülk A, Kessler A, Heck K, Hickel R, Dujic H, et al. The Detection of Dental Pathologies on Periapical Radiographs—Results from a Reliability Study. Vol. 12, Journal of Clinical Medicine. 2023.

20. Zhang CN, Zhu Y, Fan LF, Zhang X, Jiang YH, Gu YX. Intra- and inter-observer agreements in detecting peri-implant bone defects between periapical radiography and cone beam computed tomography: A clinical study. J Dent Sci [Internet]. 2021;16(3):948–56. Available from: https://www.sciencedirect.com/science/article/pii/S1991790220302361

21. Wenzel A. Radiographic modalities for diagnosis of caries in a historical perspective: from film to machine-intelligence supported systems. Dentomaxillofac Radiol. 2021 Jul;50(5):20210010.

22. Gandhi TK, Classen D, Sinsky CA, Rhew DC, Vande Garde N, Roberts A, et al. How can artificial intelligence decrease cognitive and work burden for front line practitioners? JAMIA open. 2023 Oct;6(3):ooad079.

23. van Nooijen CCA, de Koning BB, Bramer WM, Isahakyan A, Asoodar M, Kok E, et al. A Cognitive Load Theory Approach to Understanding Expert Scaffolding of Visual Problem-Solving Tasks: A Scoping Review. Educ Psychol Rev [Internet]. 2024;36(1):12. Available from: https://doi.org/10.1007/s10648-024-09848-3

24. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Futur Healthc J. 2021 Jul;8(2):e188–94.

25. Khalifa M, Albadawy M. AI in diagnostic imaging: Revolutionising accuracy and efficiency. Comput Methods Programs Biomed Updat [Internet]. 2024;5:100146. Available from: https://www.sciencedirect.com/science/article/pii/S2666990024000132

26. Mahesh Batra A, Reche A. A New Era of Dental Care: Harnessing Artificial Intelligence for Better Diagnosis and Treatment. Cureus. 2023 Nov;15(11):e49319.

27. Anichini G, Natali C, Cabitza F. Invisible to Machines: Designing AI that Supports Vision Work in Radiology. Comput Support Coop Work [Internet]. 2024; Available from: https://doi.org/10.1007/s10606-024-09491-0

28. Badal K, Lee CM, Esserman LJ. Guiding principles for the responsible development of artificial intelligence tools for healthcare. Commun Med. 2023 Apr;3(1):47.

Chapter 9

9.1 Ethics approval letter (chapters 3 and 4)

Dr Stephen Cox
Sydney Dental School; Faculty of Medicine and Health
Email: stephen.cox@sydney.edu.au

Dear Stephen,

The University of Sydney Human Research Ethics Committee (HREC) has considered your application.

I am pleased to inform you that after consideration of your response, your project has been approved.

Details of the approval are as follows:

Project No.: **2020/336**

Project Title: Authorised **Errors of Interpretation of Dental Radiographs among Dental Professionals: A Survey**

Personnel: Approval **Cox Stephen; Gao Jinlong; Hegde Shwetha; Vasa Rajesh;**

Period: **10 July 2020 to 10 July 2024**

First Annual Report Due: **10 July 2021**

Documents Approved:

Date Uploaded	Version Number	Document Name
22/06/2020	Version 3	PIS revised clean
22/06/2020	Version 1	ADA newsletter advert
22/04/2020	Version 1	email invitation template
30/01/2020	Version 1	Survey Questionnaire

Condition/s of Approval

- Research must be conducted according to the approved proposal.
- An annual progress report must be submitted to the Ethics Office on or before the anniversary of approval and on completion of the project.
- You must report as soon as practicable anything that might warrant review of ethical approval of the project including:
  - Serious or unexpected adverse events (which should be reported within 72 hours).
  - Unforeseen events that might affect continued ethical acceptability of the project.
- Any changes to the proposal must be approved prior to their implementation (except where an amendment is undertaken to eliminate *immediate* risk to participants).
- Personnel working on this project must be sufficiently qualified by education, training and experience for their role, or adequately supervised. Changes to personnel must be reported and approved.
- Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, as relevant to this project.

- Data and primary materials must be retained and stored in accordance with the relevant legislation and University guidelines.
- Ethics approval is dependent upon ongoing compliance of the research with the *National Statement on Ethical Conduct in Human Research*, the *Australian Code for the Responsible Conduct of Research*, applicable legal requirements, and with University policies, procedures and governance requirements.
- The Ethics Office may conduct audits on approved projects.
- The Chief Investigator has ultimate responsibility for the conduct of the research and is responsible for ensuring all others involved will conduct the research in accordance with the above.

This letter constitutes ethical approval only.

Please contact the Ethics Office should you require further information or clarification.

Sincerely,

[redacted]

Associate Professor Michael Skilton
Chair, Health Review Committee (Low Risk)

The University of Sydney of Sydney HRECs are constituted and operate in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2018) and the NHMRC's Australian Code for the Responsible Conduct of Research (2018)

9.2 Ethics approval letter (chapter 6)

THE UNIVERSITY OF SYDNEY

Research Integrity & Ethics Administration
HUMAN RESEARCH ETHICS COMMITTEE

Tuesday, 31 August 2021

Dr Shwetha Hegde
Sydney Dental School; Faculty of Medicine and Health
Email: shwetha.hegde@sydney.edu.au

Dear Shwetha,

The University of Sydney Human Research Ethics Committee (HREC) has considered your application.

I am pleased to inform you that after consideration of your response, your project has been approved.

Details of the approval are as follows:

Project No.:	**2021/454**
Project Title:	**Attitudes and Perceptions of Australian Dental Practitioners and Dental Students to Applications of Artificial Intelligence in Dentistry: A Survey**
Authorised Personnel:	**Hegde Shwetha; Gao Jinlong; Luu Vivian; Jordan Ashleigh; Patel Usaamah; Jeha Omar;**
Approval Period:	**31 August 2021** to **31 August 2025**
First Annual Report Due:	**31 August 2022**

Documents Approved:

Date Uploaded	Version Number	Document Name
20/08/2021	Version 1	Email to other HoS
06/08/2021	Version 2	Participant Information Form revised clean
06/08/2021	Version 2	LMS announcement revised
06/08/2021	Version 2	revised questionnaire
06/08/2021	Version 1	Flyer/ circular
06/08/2021	Version 1	LMS announcement for other universities
06/08/2021	Version 1	invitation to participate in the survey/email
03/06/2021	Version 1	AI survey advert and flyer
03/06/2021	Version 1	email template

Condition/s of Approval

- Research must be conducted according to the approved proposal.
- An annual progress report must be submitted to the Ethics Office on or before the anniversary of approval and on completion of the project.
- You must report as soon as practicable anything that might warrant review of ethical approval of the project including:
  - Serious or unexpected adverse events (which should be reported within 72 hours).
  - Unforeseen events that might affect continued ethical acceptability of the project.
- Any changes to the proposal must be approved prior to their implementation (except where an amendment is undertaken to eliminate *immediate* risk to participants).
- Personnel working on this project must be sufficiently qualified by education, training and experience for their role, or adequately supervised. Changes to personnel must be reported and approved.

- Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, as relevant to this project.
- Data and primary materials must be retained and stored in accordance with the relevant legislation and University guidelines.
- Ethics approval is dependent upon ongoing compliance of the research with the *National Statement on Ethical Conduct in Human Research*, the *Australian Code for the Responsible Conduct of Research*, applicable legal requirements, and with University policies, procedures and governance requirements.
- The Ethics Office may conduct audits on approved projects.
- The Chief Investigator has ultimate responsibility for the conduct of the research and is responsible for ensuring all others involved will conduct the research in accordance with the above.

This letter constitutes ethical approval only.

Please contact the Ethics Office should you require further information or clarification.

Sincerely,

[redacted]

Associate Professor Michael Skilton
Chair, Health Review Committee (Low Risk)

The University of Sydney of Sydney HRECs are constituted and operate in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2018) and the NHMRC's Australian Code for the Responsible Conduct of Research (2018)

9.3 Ethics approval letter (chapter 5)

Dr Hegde Shwetha
Sydney Dental School; Faculty of Medicine and Health
Email: shwetha.hegde@sydney.edu.au

Dear Dr Shwetha,

The University of Sydney Human Research Ethics Committee (HREC) has considered your application.

I am pleased to inform you that after consideration of your response, your project has been approved.

Details of the approval are as follows:

Project No.:	**2023/019**
Project Title:	**Influence of Cognitive Aids on the Accuracy of Caries Diagnosis on Bitewing Radiographs: A Comparative Study**
Authorised Personnel:	**Shwetha Hegde;**
Approval Period:	**21/04/2023** to **21/04/2027**
First Annual Report Due:	**4/21/2024**

Documents Approved:

Date Uploaded	Version Number	Document Name
05/04/2023	Version 3	Participant Information Form dentists clean
23/03/2023		Dentist advert revised clean
23/03/2023		Dentist consent form revised
23/03/2023		Student consent form revised
23/03/2023		Dentist questionnaire revised
23/03/2023	Version 2	Participant Information Form students clean
23/03/2023		Canvas announcement revised clean
15/12/2022		Post task questionnaire dental students
15/12/2022		Checklist form
15/12/2022		Figures
15/12/2022		image interpretation module TEST
15/12/2022		Caries annotation form
15/12/2022		Post task questionnaire dentists

Special Condition/s of Approval

1. Only contacts who also have publicly available contact details will be sent information about the study.
2. Approval from Head of School or Discipline is required to be kept on file to confirm that they are giving approval for staff members to be contacted about the study.

Condition/s of Approval

1. Research must be conducted according to the approved proposal.
2. An annual progress report must be submitted to the Ethics Office on or before the anniversary of approval and on completion of the project.
3. You must report as soon as practicable anything that might warrant review of ethical approval of the project including:
   - Serious or unexpected adverse events (which should be reported within 72 hours).
   - Unforeseen events that might affect continued ethical acceptability of the project.

4. Any changes to the proposal must be approved prior to their implementation (except where an amendment is undertaken to eliminate *immediate* risk to participants).

- Personnel working on this project must be sufficiently qualified by education, training and experience for their role, or adequately supervised. Changes to personnel must be reported and approved.

5. Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, as relevant to this project.

6. Data and primary materials must be retained and stored in accordance with the relevant legislation and University guidelines.

7. Ethics approval is dependent upon ongoing compliance of the research with the *National Statement on Ethical Conduct in Human Research*, the *Australian Code for the Responsible Conduct of Research*, applicable legal requirements, and with University policies, procedures and governance requirements.

8. The Ethics Office may conduct audits on approved projects.

9. The Chief Investigator has ultimate responsibility for the conduct of the research and is responsible for ensuring all others involved will conduct the research in accordance with the above.

This letter constitutes ethical approval only.

Please contact the Ethics Office should you require further information or clarification.

Sincerely,



Associate Professor Ingrid Gelissen
Acting Chair, Health Review Committee (Low Risk)

The University of Sydney of Sydney HRECs are constituted and operate in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2018) and the NHMRC's Australian Code for the Responsible Conduct of Research (2018)