Original Paper

Dental Caries Detection in Children Using Intraoral Scanners Featuring Fluorescence: Diagnostic Agreement Study

Bree Jones[1,2], MPH, PhD; Tong Chen[1,3], PhD; Stavroula Michou[4], PhD; Nicky Kilpatrick[1,5], PhD; Nigel Curtis[2,5,6], MD, MBBS, PhD; David P Burgner[2,5,6], MD, PhD; Christoph Vannahme[7], PhD; Mihiri Silva[1,2], PhD

[1]Melbourne Dental School, University of Melbourne, Carlton, Australia

[2]Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Australia

[3]Clinical Epidemiology & Biostatistics Unit, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Australia

[4]Department of Odontology, School of Health and Medical Sciences, University of Copenhagen, Nørre Allé, Denmark

[5]Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Parkville, Australia

[6]Department of Infectious Diseases, Royal Children's Hospital, Parkville, Australia

[7]3Shape TRIOS A/S, Copenhagen, Denmark

Corresponding Author:

Bree Jones, MPH, PhD
Melbourne Dental School
University of Melbourne
720 Swanston Street
Carlton 3035
Australia
Phone: 61 9341-4000
Email: bree.jones@unimelb.edu.au

Abstract

Background: Dental caries is a common chronic disease in children. Digital tools such as intraoral scanners (IOSs) may offer an efficient, scalable alternative to conventional visual examination for dental caries detection. IOSs are handheld devices that generate 3-dimensional (3D) models of the teeth and surrounding structures. Recent advances incorporating fluorescence technology into scanner hardware offer the potential for supporting dental caries detection. However, the performance of digital caries detection methods using 3D models that include both color and fluorescence in children's teeth remains unknown.

Objective: This study aimed to assess the diagnostic agreement between visual examination and on-screen assessment of 3D models generated by an IOS in color and supplemented with fluorescence for caries detection in primary teeth.

Methods: Children participating in a clinical trial (n=216, mean age 5.6, SD 0.4 y) underwent visual examination using the International Caries Detection and Assessment System (ICDAS) and intraoral scanning using the TRIOS 4 IOS. Four trained registered dental practitioners independently assessed each participant's 3D models in color and then supplemented with fluorescence, using a previously validated ICDAS index modified for on-screen assessments of 3D models. All 3D models were assessed again after 4 weeks. The time taken for intraoral scanning and on-screen assessment was recorded. Multilevel logistic regression was used to estimate and compare the likelihood of detecting caries between methods, and Bland-Altman plots were used to visualize agreement. Analyses were performed at the initial (ICDAS ≥01), moderate (ICDAS ≥03), and extensive (ICDAS ≥05) dental caries thresholds. Intraclass correlation coefficient (ICC) estimated method agreement and examiner reliability.

Results: Of 219 children enrolled, 216 completed dental assessments. A total of 9470 visible primary tooth surfaces were included in the analysis. The average time taken for on-screen assessment of each 3D model (color with fluorescence) was 3.5 (SD 2.3) minutes. The likelihood of detecting caries using color assessment of 3D models was similar to visual examination at all disease thresholds: initial (odds ratio [OR] 1.1, 95% CI 1.0-1.3), moderate (OR 0.9, 95% CI 0.7-1.1), and extensive (OR 1.0, 95% CI 0.7-1.3). When color assessments were supplemented with fluorescence, the likelihood of detecting caries was 30% higher at the initial threshold relative to visual examination (OR 1.3, 95% CI 1.1-1.5) and similar at the moderate (OR 0.9, 95% CI 0.7-1.1) and extensive thresholds (OR 0.9, 95% CI 0.7-1.3). Bland-Altman plots showed a high level of agreement at both moderate and extensive thresholds. Agreement between methods was high (ICC 0.9, 95% CI 0.9-1.0). Intra- and inter-examiner reliability using intraoral scans ranged from good to excellent (ICC 0.8-1.0).

Conclusions: On-screen assessment of 3D models in color demonstrated the highest agreement with visual examination for caries detection across all disease thresholds.

Introduction

Dental caries is a chronic, multifactorial disease influenced by environmental and behavioral factors [1]. It is estimated to impact more than half a billion children globally, representing a major public health concern [2]. Developing efficient and accurate tools to collect data on dental caries can facilitate the inclusion of oral health outcomes in epidemiological research, improving understanding of contemporary risk factors, disease distribution, and the burden of disease in children.

The clinical presentation of dental caries spans a spectrum from early to advanced disease. In its earliest stage, mineral loss causes subsurface changes that alter enamel porosity, and subsequently its optical properties. At this point, the enamel surface remains intact, and the process can be reversed or arrested through remineralization strategies [3]. If mineral loss continues, tooth structure becomes compromised, potentially leading to cavity formation. These later stages signify irreversible damage to the tooth and may eventually result in pain and tooth loss if left untreated [3].

Conventional assessment of dental caries in epidemiological research involves a visual examination by a qualified dental practitioner using a validated scoring index. Implementing visual examinations across large or multisite studies requires substantial investment in examiner training and calibration to ensure consistency, as interexaminer variability can impact study validity [4]. The financial and logistical challenges may limit the use of visual examinations, with some studies instead relying on self-reported caries status. Self-reporting methods have been criticized for their lack of validity and their tendency to underestimate disease burden in children [5].

Intraoral scanners (IOSs) are hand-held devices that generate 3-dimensional (3D) digital scans of the teeth and surrounding structures (herein 3D models). It has been suggested that 3D models can be used to perform digital dental examinations [6]. The various commercially available devices differ in their hardware and software for image acquisition and reconstruction. Manufacturers include, but are not limited to, Align Technologies (San Jose, California, USA), Medit Corp (Seoul, South Korea), Dentsply Sirona (Charlotte, NC, USA), Planmeca (Helsinki, Finland), and 3/Shape (Copenhagen, Denmark).

The TRIOS 4 and TRIOS 5 scanners (3/Shape, Copenhagen, Denmark) incorporate light-induced fluorescence technology into their hardware to aid dental caries detection. The devices use an LED light source (approximately 415 nm) and a long-pass filter to capture emitted red and green fluorescence [7]. When healthy teeth are exposed to blue light, they become excited to emit green fluorescence [8]. Areas affected by dental caries have reduced green fluorescence intensity, as mineral loss alters the tooth's composition and its light-scattering properties[9] At these wavelengths, some bacterial metabolites associated with dental caries emit red fluorescence. Using the scanner software, green and red fluorescence can be visualized simultaneously as an overlay on the 3D model. Fluorescence characteristics can be interpreted by a dental practitioner to support detection of subtle early lesions which may be difficult to visualize when viewing scans in color alone.

The International Caries Detection and Assessment System is a widely accepted scoring index for detecting and classifying dental caries severity [10]. To reduce subjectivity, a modified version incorporating fluorescence features has been developed [11] and validated [12]. In adult permanent teeth, an in vitro study investigating on-screen assessment using this index found no difference in diagnostic accuracy between on-screen assessment and conventional visual examination for occlusal tooth surfaces [13]. However, the real-world diagnostic performance of an on-screen caries assessment for primary teeth and other surface types (such as visible smooth surfaces) is unknown.

Clinically, enamel defects represent another prevalent condition that can resemble dental caries, as affected areas often display similar visual characteristics. There is also overlap in the indices used to record and quantify both conditions [14]. Yet, most studies evaluating caries detection methods typically exclude teeth with enamel defects, leaving their influence on diagnostic performance uncertain.

This study, therefore, aimed to determine the in vivo diagnostic agreement between visual examination and on-screen assessment of 3D models in approximate natural colors with and without fluorescence for detecting and classifying caries in primary teeth. The secondary aims included exploring the impact of disease threshold, enamel defects, and surface type on agreement estimates.

Methods

Study Design

A cross-sectional diagnostic agreement study was undertaken (Figure 1), nested within a longitudinal randomized controlled trial. An agreement design was selected as the visual examination comparator was not considered a gold standard reference in the study setting [15]. This study followed the Guidelines for Reporting Reliability and Agreement Studies [16] (Table SA-1 in Multimedia Appendix 1) and was

conducted per the applicable sections of the Standards for Reporting of Diagnostic Accuracy Studies [17]. The protocol was prospectively published (International Registered Report Identifier DERR1-10.2196/51578) [15] and registered with the Australian and New Zealand Clinical Trials Registry (ID ACTRN12622001237774).

Figure 1. Study design schema: study design schema illustrating the cross-sectional diagnostic agreement study nested within the MIS BAIR randomized controlled trial. The study compared the diagnostic agreement for dental caries detection and classification between visual examination and 2 on-screen assessment modalities (3D color models and 3D color with fluorescence models). Four trained examiners independently completed on-screen assessments at 2 time points (>4 wk apart). Visual examination outcomes served as the comparator. This image is created using BioRender [18].
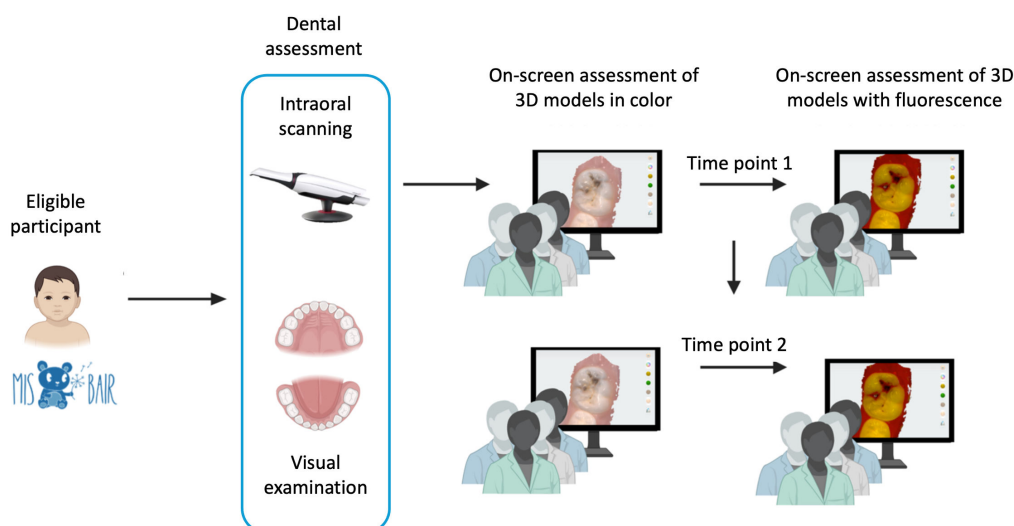
Study Setting and Participants

The subjects were participants in the Melbourne Infant Study: BCG for Allergy and Infection Reduction (MIS BAIR), a phase III multicenter, randomized controlled trial [19]. The trial aimed to investigate the impact of neonatal BCG vaccination on infection and allergy outcomes in early childhood. Eligible English-speaking mothers in Victoria (2013-2016) consented for their newborns to join MIS BAIR. Infants with contraindications to live vaccines at birth were excluded, yielding 1272 participants. Clinical assessments were conducted at 1 year and 5 years.

Trial participants received monthly newsletters and accessed study updates via the MIS BAIR website, where families were informed of an optional dental assessment (which included intraoral scanning) at the 5-year-old follow-up. The dental component commenced in January 2021, with 536 participants remaining in the trial. Despite COVID-19 disruptions, visits continued under special approval, and 219 participants had consented to dental assessments by March 2022. All dental assessments were performed at Murdoch Children's Research Institute (Melbourne, Australia).

All primary coronal tooth surfaces were eligible for inclusion in the agreement study if they were examined during the clinical visual examination and were visible for on-screen assessment on the corresponding intraoral scan. Tooth surfaces were excluded if they were located on proximal aspects, restored, sealed, had missing or insufficient scan data, or were not visible due to other reasons.

Dental Assessment Procedure

Visual examination was performed with participants in a semisupine position, under illumination, using a dental mirror and probe with no magnification. Two trained and calibrated dental practitioners examined all teeth after removing gross plaque deposits using Puritan swabs and drying teeth with cotton wool rolls and gauze. Training procedures are described elsewhere [15]. Universal infection control standards and additional COVID-19 precautions were adhered to for all examinations.

All tooth surfaces were examined following a standard operating procedure. Any reasons for noncompletion of any parts of the dental examination were recorded. Dental caries was recorded using a merged International Caries Detection and Assessment System (ICDAS) criteria [11], and enamel defects were recorded using a modified version of the European Academy of Paediatric Dentistry index [20]. If there were multiple lesions present on the one tooth surface, the most severe of the 2 clinical presentations was recorded.

Teeth were scanned with the TRIOS 4 IOS immediately following the visual examination. The laptop was set up on a table across from the participant, where they were able to view scan acquisition during scanning (Figure 2). A scan was deemed adequate when sufficient color and fluorescence data were captured. Incomplete scans were documented, with common causes, including scanner tip access, gag reflex, or limited cooperation with scanning due to behavioral or developmental factors. Each scan was saved under a unique identifier using the TRIOS Dental Desktop software with

the TRIOS module. The scanning time was calculated using the TRIOS software. The IOSs were regularly calibrated according to the manufacturer's procedures. The manufacturer scanning strategies were followed, and external light was minimized during scanning.

Figure 2. Intraoral scanning procedure. A participant undergoing intraoral scanning as part of this research. The participant can observe the intraoral scan being constructed in real time during the procedure.


On-Screen Assessment Procedure

On-screen assessments of 3D models commenced in November 2022, 8 months after the dental assessments. Examiner training and calibration processes for on-screen assessments are detailed in the study protocol [15] and Table SA-2 in Multimedia Appendix 1). Assessments were performed on 15-inch monitors using Windows 11 in rooms with natural lighting. A custom noncommercial software (developed by 3Shape) was used to efficiently retrieve and view multiple postprocessed 3D models at resolution quality equivalent to TRIOS Dental Desktop. A systematic visual examination protocol guided all on-screen assessments. Four trained dental examiners, blinded to the clinical data and each other's ratings, independently evaluated each participant's models in 2 stages: first assessing surface geometry in approximate natural colors and then supplemented with fluorescence data (Figure 3). The presence and extent of dental carious lesions were recorded using a validated, modified ICDAS index incorporating fluorescence descriptors (Figure SA-1, Multimedia Appendix 1) [11-13,15]. The models were assessed at time point 1. All examiners repeated the on-screen assessment protocol >4 weeks later (time point 2).

All data were recorded and managed using a secure, web-based app, the REDCap (Research Electronic Data Capture) tool [21].

Figure 3. 3D models generated by an intraoral scanner. Left: Approximate natural color textures. Right: With the addition of fluorescence texture.
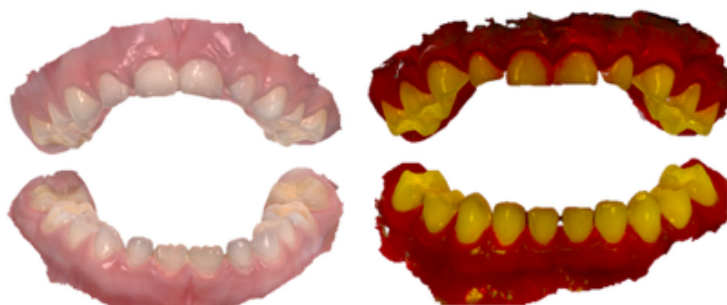
Outcome Measures

The primary outcome was dental caries at a tooth surface level. Three binary outcome variables were derived to represent clinically relevant disease thresholds for dental caries. Specifically, caries data were dichotomized at initial (ICDAS ≥01), moderate (ICDAS ≥03), and extensive disease thresholds (ICDAS ≥05). All other variable definitions are described in Table SA-3 (Multimedia Appendix 1).

Sample Size

The sample was drawn from all 219 eligible children who participated in the MIS BAIR dental assessment at the 5-year-old follow-up visit.

Statistical Analysis

Evaluating tooth surface-level outcomes in dental research introduces complexities to the statistical analysis and limits the use of traditional methods for examining diagnostic agreement. Whole mouth dental data have multiple levels of clustering (surfaces nested within a tooth and teeth nested within an individual), and the high prevalence of healthy tooth surfaces results in a zero-inflated data distribution.

First, to account for the clustering, this study used multilevel logistic regression to estimate the likelihood of caries detection at the tooth surface level. The likelihood of detecting dental caries between visual examination and on-screen assessment methods (with and without fluorescence) was estimated by including an indicator variable for the caries detection method as a fixed effect. Additionally, the 2 on-screen assessment methods were compared by fitting a separate model that included method, rater, and time (time point 1 or time point 2) as fixed effects. A sensitivity analysis was performed by adjusting the model for enamel defects, and an additional sensitivity analysis by surface type (smooth surfaces vs occlusal surfaces) investigated if estimates differed based on surface location.

Second, the agreement between methods for the total number of caries detected for each participant was examined using the Bland-Altman plot [22]. The Bland-Altman plot compared the differences in the total number of carious surfaces detected between methods (y-axis) against the mean of the 2 methods (x-axis) for each subject. Due to the significant overlap of points, the size of the points in the plot was adjusted to reflect the frequency of the observations. Visual examination was compared to on-screen assessment with and without fluorescence at each diagnostic threshold (initial, moderate, and extensive). This comparison was repeated for each surface type (occlusal and smooth surfaces). A best-fit regression line of the differences in methods against the mean was plotted to identify trends in differences across the range of averages.

Finally, the intraclass correlation coefficient (ICC) was calculated to estimate the concordance between visual examination and on-screen assessment methods (with and without fluorescence) for total carious surfaces detected by each method per participant. The ICC for count data was estimated from generalized linear mixed models with zero-inflated Poisson distribution to account for the excess zero observations [23]. For the primary method comparison, given the multiple examiner observations for the on-screen assessment methods, the median number of surfaces with carious lesions detected among examiners for each participant was used to compare to the visual examination. Additionally, the ICC was calculated for each combination of raters to determine the concordance (interrater reliability) for on-screen assessments (with and without fluorescence) and the repeatability (intrarater reliability) for the same rater between time point 1 and time point 2 for on-screen assessments. All analyses were performed in STATA version 18 (StataCorp) [24] and R (R Core Team) [25].

Ethical Considerations

This research was conducted with the approval of the Royal Children's Hospital Human Research Ethics Committee (RCH HREC No.88321). Written parental or guardian consent was obtained for each participant to allow inclusion in the study and the use of dental data for validating and advancing intraoral scanning technology for oral health assessment. Children received a AUD $20 gift card (US $15) as remuneration for their participation. All participant records were assigned unique identifiers, ensuring that individual identities were not identifiable by the dental team. The study was conducted according to the principles outlined in the Declaration of Helsinki and the ICMJE recommendations. The study protocol is registered with the Australian New Zealand Clinical Trials Registry (registration number: ACTRN12622001237774p). The study protocol was published a priori with the International Registered Report Identifier DERR1-10.2196/51578. The MIS BAIR randomized controlled trial had ethical and governance approval from Mercy Health Human Research Ethics Committee (HREC; No. R12-28) and Royal Children's Hospital HREC (No. 33025).

Results

Study Flow

Of the 219 children who consented to dental assessments, 3 declined dental examinations or intraoral scanning at the time of the study visit, leaving a study population of 216. A total of 18,887 tooth surfaces were eligible for inclusion. Permanent tooth surfaces (n=750), proximal tooth surfaces (n=8209), restored or sealed tooth surfaces (n=120), and surfaces that were not visible or that could not be examined due to incomplete or missing scan data (n=338) were excluded. The remaining 9470 (50%) tooth surfaces were included in the analysis (Figure 4).

Figure 4. Study flow diagram. Flow diagram outlining patient eligibility and exclusion criteria (N=number of patients, n=number of tooth surfaces).
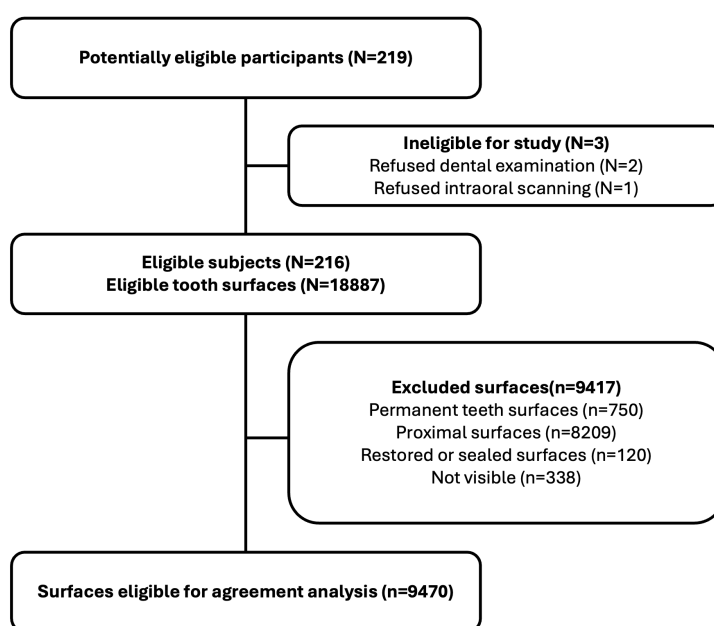
Sample Characteristics

Table 1 presents the characteristics of the sample. Participants' average age was 5.6 (SD 0.4) years, with 49% males and 51% females. Most children had mothers of Anglo-Celtic ancestry (n=124, 57%), followed by European (n=38, 18%) and Asian (n=33, 15%). Most children resided in the state's most advantaged areas (n=138, 64%). On the basis of visual examination data, the prevalence of dental caries at the initial disease threshold was 38%, and enamel defects were 18%.

Table 1. Sample characteristics (N=216).

Participant characteristics	Values
Age (y), mean (SD)	5.6 (0.4)
Sex, n (%)	
Male	106 (49.1)
Female	110 (50.9)
Ethnicity, n (%)	
Anglo-Celtic	124 (57.4)
European	38 (17.6)
Asian	33 (15.3)
Middle East	4 (1.9)
African	4 (1.9)
Aboriginal/Torres Strait Islander	2 (0.9)
Other	11 (5.1)
SEIFA[a], n (%)	
Quintile 1	5 (2.3)
Quintile 2	31 (14.4)
Quintile 3	42 (19.4)
Quintile 4	49 (22.7)
Quintile 5	89 (41.2)
Prevalence of caries	82 (37.9)
Prevalence of enamel defects	39 (18.1)

[a]Socio-economic Indexes for Areas (SEIFA) ranks areas in Australia according to relative socioeconomic disadvantage. SEIFA was divided into quintiles ranging from the most disadvantaged (quintile 1) to the least disadvantaged (quintile 5).

Examiner Characteristics

Four registered dental practitioners undertook on-screen assessments for all participants at 2 time points. They were recruited via convenience sampling. The examiners had clinical experience ranging from 5 to 15 years, including conducting epidemiological assessments (Table SA-2 in Multimedia Appendix 1).

Time Taken for Scanning and On-Screen Assessments

The mean time to perform intraoral scanning of the lower arch was 46 (SD 8) seconds and for the upper arch was 58 (SD 15) seconds. The mean time taken for the on-screen assessment of each 3D model in color supplemented with fluorescence was 3.5 (SD 2.3) minutes.

Surface-Level Agreement

The total number of carious tooth surfaces reported for each method at each diagnostic threshold and odds ratios (ORs) are presented in Table 2, Table 3, and Table 4 (Raw data stratified by rater and surface type are available in Table SA-4 and Table SA-5 Multimedia Appendix 1). The likelihood of detecting caries using on-screen assessment of 3D models in color was similar to that of visual examination at the initial (odds ratio [OR] 1.1, 95% CI 1.0-1.3), moderate (OR 0.9, 95% CI 0.7-1.1), and extensive (OR 1.0, 95% CI 0.7-1.3) disease thresholds. When 3D models were supplemented with fluorescence, the likelihood of detecting caries was 30% higher at the initial threshold relative to visual examination (OR 1.3, 95% CI 1.1-1.5) and similar at the moderate (OR 0.9, 95% CI 0.7-1.1) and extensive thresholds (OR 0.9, 95% CI 0.7-1.3). The sensitivity analysis by surface type is also presented in Table SA-6 in Multimedia Appendix 1. The likelihood of caries detection using on-screen assessment with and without fluorescence was comparable to visual examination for smooth surfaces across all disease thresholds. For occlusal surfaces, the likelihood of reporting caries with on-screen assessment methods was 1.5- to 2-fold higher at the initial disease threshold both with (OR 2.0, 95% CI 1.6-2.4) and without fluorescence (OR 1.5, 95% CI 1.2-1.9). The likelihood of reporting at the moderate and extensive thresholds on occlusal surfaces was similar for on-screen assessments and visual examinations. Comparing on-screen assessment methods reveals that adding fluorescence increases the likelihood of caries detection, and this likelihood is also influenced by the rater scoring the assessment. Accounting for enamel defects as an additional covariate in the analysis resulted in slightly lower OR estimates for both methods across all disease thresholds.

Table 2. Agreement estimates for visual examination (VE) versus on-screen assessment (OSA) for all included tooth surfaces (n=9470).

Disease threshold	VE		OSA		Main analysis[a]		Sensitivity analysis	
	No	Yes	No	Yes	OR (95% CI)	P value	OR (95% CI)	P value
Initial threshold	9182	288	9154	316	1.1 (1.0-1.3)	0.08	1.1 (1.0-1.3)	0.03
Moderate threshold	9384	86	9395	75	0.9 (0.7-1.1)	0.19	0.8 (0.7-1.1)	0.17
Extensive threshold	9422	48	9424	46	1.0 (0.7-1.3)	0.78	1.0 (0.7-1.3)	0.76

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.

Table 3. Agreement estimates for visual examination (VE) versus on-screen assessment with fluorescence (OSA-FLU) for all included tooth surfaces (n=9470).

Disease threshold	VE		OSA- FLU		Main analysis[a]		Sensitivity analysis	
	No	Yes	No	Yes	OR (95% CI)	P value	OR (95% CI)	P value
Initial threshold	9182	288	9125	345	1.3 (1.1-1.5)	<.001	1.2 (1.1-1.4)	<.001
Moderate threshold	9384	86	9394	76	0.9 (0.7-1.1)	0.24	0.9 (0.7-1.0)	0.21
Extensive threshold	9422	48	9380	45	0.9 (0.7-1.3)	0.61	0.9 (0.7-1.3)	0.59

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.

Table 4. Agreement estimates for on-screen assessment (OSA) versus on-screen assessment with fluorescence (OSA-FLU) for all included tooth surfaces (n=9470).

Disease threshold	OSA No	OSA Yes	OSA- FLU No	OSA- FLU Yes	Main analysis[a] OR (95% CI)	P value	Sensitivity analysis OR (95% CI)	P value
Initial threshold	9154	316	9125	345	1.1 (1.1-1.2)	<.001	1.1 (1.1-1.2)[b]	<.001
Moderate threshold	9395	75	9394	76	1.0 (0.9-1.2)	0.77	1.0 (0.9-1.2)[b]	0.77
Extensive threshold	9424	46	9380	45	1.0 (0.8-1.1)	0.62	1.0 (0.8-1.1)	0.62

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.
[b]The inclusion of a rater variable in this model had a significant effect.

Person-Level Agreement

The Bland-Altman plots indicate the absolute method agreement (Figure 5). For on-screen assessment of 3D models without (OSA) and with fluorescence (OSA-FLU) compared to visual examination, the absolute agreement was lower at the initial disease threshold and higher at the moderate and extensive disease thresholds. At the initial disease threshold, there were more negative differences compared to positive differences, where on-screen assessment recorded more carious surfaces relative to the visual examination (Figure 4A and D). At the moderate and extensive thresholds, the frequencies of positive and negative differences were comparable for both color-only and fluorescence-assisted assessments (Figure 4B, C, E, and F). The slope and intersection of the regression line, particularly at the initial caries thresholds, suggest that when average total carious surfaces were low, there were negative differences between methods. As the average total carious surfaces increased, the positive difference between methods increased (where visual examination recorded more carious surfaces relative to on-screen assessment). The subanalysis based on surface type shows similar trends (Figure SA-2 in Multimedia Appendix 1).

Figure 5. Bland-Altman plots of the differences between the methods for all tooth surfaces. Bland-Altman plots comparing the differences in the number of carious lesions detected for each pair of methods against the average of the 2 methods across the initial, moderate, and extensive disease thresholds. The regression line of the differences in methods against the mean is represented by a red dash. The size of the data points is proportional to the number of observations. (A) Initial threshold, OSA versus VE: absolute agreement n=67 (31%), negative differences n=108 (50%), positive differences n=41 (19%). (B) Moderate threshold, OSA versus VE: absolute agreement n=170 (79%), negative differences n=26 (12%), and positive difference n=20 (9%). (C) Extensive threshold, OSA versus VE: absolute agreement n=192 (89%), negative differences n=14 (6%), positive difference n=10 (5%). (D) Initial threshold, OSA-FLU versus VE: absolute agreement n=64 (30%), negative differences n=112 (52%), positive differences n=40 (18%). (E) Moderate threshold, OSA-FLU versus VE: absolute agreement n=167 (77%), negative differences n=28 (13%), positive differences n=21 (10%). (F) Extensive threshold, OSA-FLU versus VE: absolute agreement n=192 (89%), negative differences n=14 (6%), positive difference n=10 (5%). VE: visual examination; OSA: on-screen assessment in color only; OSA FLU: on-screen assessment with fluorescence.
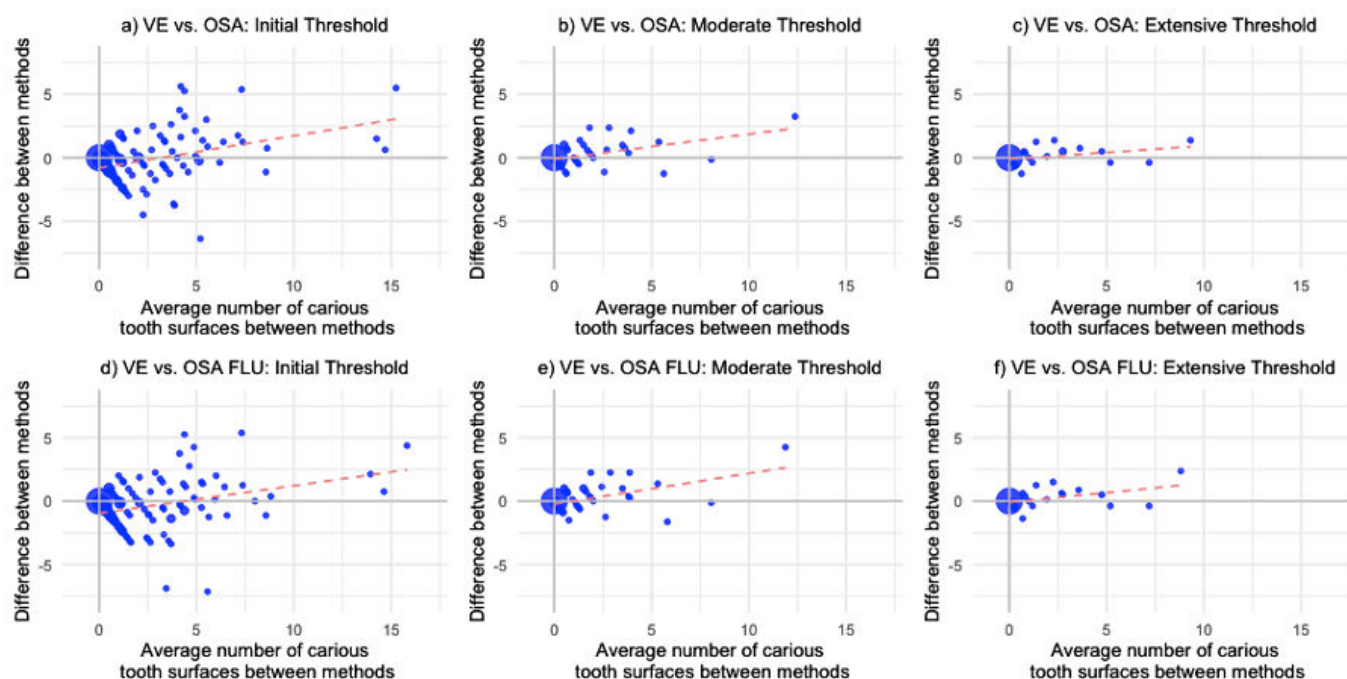
Table 5 presents the ICC values as a measure of method agreement and examiner reliability. On-screen assessment in color (ICC 0.9, 95% CI 0.9-0.9) and with fluorescence (ICC 0.9, 95% CI 0.8-0.9) showed good to excellent agreement with visual examination. For all examiners, the intrarater agreement of ICC of approximately >0.9 indicated a very high degree of scoring consistency within raters for on-screen assessment methods. The interrater reliability was also high between all raters (>0.9), indicating high-scoring similarity between raters. The narrow CIs further support the reliability of these estimates.

Table 5. Intraclass correlation coefficients (ICC). The ICC was used to estimate the absolute agreement between methods, intraexaminer reliability for on-screen assessment methods, and interrater agreement reliability for on-screen assessment methods.

	On-screen assessment with color, ICC (95% CI)	On-screen assessment with color and fluorescence, ICC (95% CI)
Agreement between methods		
Visual examination	0.9 (0.9-1.0)	0.9 (0.8-1.0)
Intrarater reliability		
Rater 1	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 2	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 4	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Interrater reliability		
Rater 1 versus rater 2	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 1 versus rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 1 versus rater 4	0.9 (0.8-0.9)	0.9 (0.8-0.9)
Rater 2 versus rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 2 versus rater 4	0.9 (0.8-0.9)	0.9 (0.9-0.9)
Rater 3 versus rater 4	0.9 (0.8-0.9)	0.9 (0.8-0.9)

Discussion

Principal Findings

This study is the first to describe the real-world performance of IOSs using light-induced fluorescence for digital caries assessments of primary teeth in children. The findings suggest that the on-screen assessment of 3D models in color supplemented with fluorescence had a high level of agreement with visual examination for detecting primary coronal carious lesions on smooth and occlusal surface types at the moderate and extensive disease thresholds. At the initial disease thresholds, on-screen assessments resulted in a higher reporting of occlusal caries relative to visual examination, and supplementing on-screen color assessments with fluorescence increased the likelihood of dental caries detection compared to visual examination. The high ICC values provide further evidence of method agreement and indicate that the on-screen assessment criteria can be applied consistently and are reproducible.

Our findings are consistent with data from previous studies in adults. Porumb et al [26] observed high rater reliability and agreement between visual examination using ICDAS and on-screen assessment of 3D color models in adults when experienced examiners were used. Building on the work of a small pilot study that suggested remote assessment of 3D models in adults had good accuracy for cavity detection [6], our analysis shows that when applying a validated scoring index with trained examiners, assessment of intraoral scans has substantial agreement with visual examination.

Our study results also corroborate the findings of several studies investigating on-screen assessments for permanent occlusal surfaces [12,13], showing that on-screen assessments had similar diagnostic performance to visual examinations. However, in contrast to these studies, our findings at the initial disease threshold, showed adding fluorescence increased the likelihood of reporting dental caries for occlusal surfaces.

The results of our study align with findings from prior research using intraoral cameras for dental caries data collection in pediatric epidemiological settings. In 5-year-old children, Boye et al [27,28] found the on-screen assessment of 2D images produced with intraoral cameras was equivalent to visual examination, although they used advanced disease thresholds (cavity involving dentin) for caries reporting. A study by Mehta et al [29] involving Australian 5-year-old children found that the on-screen assessment of 2D images resulted in significantly greater reporting of dental caries compared to visual examination when initial disease thresholds were applied, which the authors attributed to the improved lighting and examination conditions when performing on-screen assessments.

In the absence of a gold standard reference, it is not possible to determine whether the observed differences between methods reflect greater accuracy of one method over the other. The higher likelihood of reporting dental caries when on-screen assessment was supplemented with fluorescence could be attributable to several factors. It is plausible that on-screen assessment is more sensitive to early mineral loss than visual examination due to enhanced detection

capabilities of light-induced fluorescence [30]. However, the apparent increase in detection may also reflect a reduction in specificity, as red fluorescence changes may be attributed to the presence of plaque leading to false-positive indications, particularly for occlusal surfaces, which often have plaque retentive pits and fissures [31]. The latter is consistent with findings investigating caries detection using 2D image-based ICDAS methods, which have reported lower specificity at the initial disease threshold [32]. The inherent subjectivity of early occlusal caries detection may also have contributed to the higher disagreements observed in our study at the initial disease threshold. Additionally, the creamy-colored demarcations associated with hypomineralized enamel could be misinterpreted as caries during the on-screen assessment, consistent with the reduced odds ratio estimates when enamel defects were included as a covariate in our sensitivity analysis.

Strengths

The use of intraoral scanning for data collection in research offers several advantages. First, it enables examiner blinding to participant characteristics and allows repeat examinations, both of which improve internal consistency and reliability. Intraoral scanning may also facilitate more comprehensive dental assessments for children who can tolerate only brief or limited inspections. The resulting 3D models can be enlarged, rotated, and reviewed for extended periods, providing access to areas that may be difficult to view during visual examination alone. Furthermore, 3D models can serve as effective communication tools. A randomized controlled trial reported that 3D models as a visual aid significantly improved parents' understanding of their children's dental findings and treatment plans in pediatric dentistry compared to verbal explanations alone [33]. Finally, the data acquisition and on-screen assessment times achieved in this study were shorter than those reported in previous research using intraoral cameras for epidemiological research, which ranged from 6- to 8.8 minutes for data acquisition per participant [28].

Limitations

We acknowledge several limitations. First, the time taken to undertake each dental examination was not recorded for all participants. Future studies should capture this, as it helps contextualize the clinical relevance. Second, the examination site lacked compressed air, and teeth were not professionally cleaned before examination and scanning. Third, unlike on-screen assessment, we did not undertake repeated visual examinations by multiple examiners. This is a potential source of bias, as it impacts the reliability of the visual examination as the comparator. Fourth, the highly selected sample potentially limits the broader generalizability of the findings. Additionally, a larger sample would increase the precision around the odds ratio estimates. Fifth, the study used an agreement design, which, although appropriate for this investigation, limits interpretation to relative rather than absolute diagnostic performance, as no gold standard reference was available. Sixth, dental caries vary in clinical presentation on proximal surfaces and around restorations or

sealants, which were excluded from the analysis due to a lack of radiographic imaging in the study design. Given the age and disease experience of the cohort, it is not known how missing teeth, widened interdental spaces, recession, or heavily restored teeth may impact on-screen assessment performance. Further validation research in different population groups with varying levels of disease is warranted. Finally, the nonnormal distribution of the dataset precluded the calculation of Bland-Altman Limits of Agreement, which define an acceptable disagreement range among methods. Data transformations to address this lacked clinical interpretability. Clinically acceptable limits of agreement were therefore not defined, as establishing these would require a consensus-based approach, which was beyond the scope of this study.

Future Research Directions

A number of practical considerations must be addressed before intraoral scanning can be implemented at scale in epidemiological or public health programs. These include assessing the resource requirements, examiner training needs, and workflow integration within existing service models. Feasibility will also depend on whether the technology is valid and reliable for the specific disease outcomes being evaluated, as different conditions may require tailored scanning protocols or interpretation methods. Additionally, the acceptability, cost-effectiveness, and environmental impact of integrating intraoral scanning into outreach, screening, and teledentistry programs should be evaluated. Research exploring the potential for trained nondental personnel to collect scan data could further support the scalability and sustainability of this approach. Furthermore, investigation into methods that automate caries detection using intraoral scan data is necessary to reduce the time and potential costs associated with using registered dental practitioners for on-screen assessments. This could be achieved through the validation of existing tools [7,34,35] or the development of new artificial intelligence methods [36]. In epidemiological settings, digital data collection combined with automation may potentially improve the frequency of population disease surveillance programs, where data-driven insights into regional oral health trends could be used to inform resource allocation and health promotion efforts.

Conclusions

This study found that the on-screen assessment of 3D models supplemented with fluorescence substantially agreed with visual examination for caries detection at the moderate and extensive disease thresholds. Whether the on-screen assessment of 3D models generated with IOSs can be used interchangeably with visual examination for caries detection in children is context dependent and likely influenced by the caries threshold used and the tooth surfaces examined. These findings suggest that digital assessments may provide a feasible alternative to traditional in-person examinations for research, longitudinal monitoring, and population-level surveillance, especially where examiner access is limited or standardization across examiners and sites is required.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to consent not obtained from participants for public sharing of data but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BJ (lead), MS
Data curation: BJ (lead), MS (supporting), NC (supporting)
Formal analysis: BJ (lead) TC (supporting)
Funding acquisition: CV,
Investigation: BJ (lead)
Methodology: BJ (lead), TC (equal)
Project administration: BJ, MS
Resources: SM, CV
Supervision: MS (lead) NK, DB, NC
Validation: BJ Visualization: BJ, TC
Writing – original draft: BJ (lead)
Writing – review & editing: BJ, TC, SM, NK, DB, NC, CV

Conflicts of Interest

BJ reports the company 3/Shape TRIOS A/S Denmark funded her PhD stipend through an industrial agreement with Murdoch Children's Research Institute. CV and SM were employed at 3Shape when this research was undertaken. This project was funded through a Research Collaboration between 3Shape TRIOS A/S Denmark and the Murdoch Children's Research Institute, in which MS is the lead investigator. Other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Multimedia Appendix 1

Additional tables and figures provided to support the findings reported in this manuscript.
[DOCX File (Microsoft Word File), 656 KB-Multimedia Appendix 1]

References

1. Pitts NB, Zero DT, Marsh PD, et al. Dental caries. Nat Rev Dis Primers. May 25, 2017;3(1):17030. [doi: 10.1038/nrdp.2017.30] [Medline: 28540937]
2. Global Oral Health Status Report: Towards Universal Health Coverage for Oral Health by 2030. WHO; 2022. URL: https://www.who.int/publications/i/item/9789240061484 [Accessed 2025-11-25] ISBN: 978-92-4-006148-4
3. Frencken JE. Dental caries and caries epidemiology. In: Evidence-Based Caries Prevention. Springer, Cham; 2025:1-21. [doi: 10.1007/978-3-031-79152-9_1]
4. Nelson S, Eggertsson H, Powell B, et al. Dental examiners consistency in applying the ICDAS criteria for a caries prevention community trial. Community Dent Health. Sep 2011;28(3):238-242. [Medline: 21916361]
5. Silva AER, Menezes AMB, Assunção MCF, et al. Validation of self-reported information on dental caries in a birth cohort at 18 years of age. PLoS ONE. 2014;9(9):e106382. [doi: 10.1371/journal.pone.0106382] [Medline: 25202969]
6. Steinmeier S, Wiedemeier D, Hämmerle CHF, Mühlemann S. Accuracy of remote diagnoses using intraoral scans captured in approximate true color: a pilot and validation study in teledentistry. BMC Oral Health. Sep 25, 2020;20(1):266. [doi: 10.1186/s12903-020-01255-8] [Medline: 32977794]
7. Michou S, Benetti AR, Vannahme C, Hermannsson PG, Bakhshandeh A, Ekstrand KR. Development of a fluorescence-based caries scoring system for an intraoral scanner: an in vitro study. Caries Res. 2020;54(4):324-335. [doi: 10.1159/000509925] [Medline: 33053552]

8. Karlsson L. Caries detection methods based on changes in optical properties between healthy and carious tissue. Int J Dent. 2010;2010:270729. [doi: 10.1155/2010/270729] [Medline: 20454579]

9. de Josselin de Jong E, Sundström F, Westerling H, Tranaeus S, ten Bosch JJ, Angmar-Månsson B. A new method for in vivo quantification of changes in initial enamel caries with laser fluorescence. Caries Res. 1995;29(1):2-7. [doi: 10.1159/000262032] [Medline: 7867045]

10. Criteria manual – international caries detection and assessment system (ICDAS II). International Caries Detection and Assessment System (ICDAS) Coordinating Committee; 2005. URL: https://www.iccms-web.com/uploads/asset/592848be55d87564970232.pdf [Accessed 2025-11-25]

11. Ferreira Zandona A, Santiago E, Eckert G, Fontana M, Ando M, Zero DT. Use of ICDAS combined with fluorescence for caries detection. Caries Res. 2010;44(3):317-322. [doi: 10.1159/000317294] [Medline: 20588022]

12. Sá G, Michou S, Bönecker M, Mendes F, Amarante B, Ekstrand K. Diagnostic validity of ICDAS clinical criteria on digital 3D models. J Dent. Oct 2024;149:105274. [doi: 10.1016/j.jdent.2024.105274] [Medline: 39084547]

13. Ntovas P, Michou S, Benetti AR, et al. Occlusal caries detection on 3D models obtained with an intraoral scanner. A validation study. J Dent. Apr 2023;131:104457. [doi: 10.1016/j.jdent.2023.104457] [Medline: 36858167]

14. Lygidakis NA, Garot E, Somani C, Taylor GD, Rouas P, Wong FSL. Best clinical practice guidance for clinicians dealing with children presenting with molar-incisor-hypomineralisation (MIH): an updated European Academy of Paediatric Dentistry policy document. Eur Arch Paediatr Dent. Feb 2022;23(1):3-21. [doi: 10.1007/s40368-021-00668-5] [Medline: 34669177]

15. Jones B, Michou S, Chen T, et al. Caries detection in primary teeth using intraoral scanners featuring fluorescence: protocol for a diagnostic agreement study. JMIR Res Protoc. Dec 14, 2023;12:e51578. [doi: 10.2196/51578] [Medline: 38096003]

16. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. Jan 2011;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]

17. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. Nov 14, 2016;6(11):e012799. [doi: 10.1136/bmjopen-2016-012799] [Medline: 28137831]

18. BioRender. URL: https://BioRender.com [Accessed 2025-12-04]

19. Messina NL, Gardiner K, Donath S, et al. Study protocol for the Melbourne Infant Study: BCG for Allergy and Infection Reduction (MIS BAIR), a randomised controlled trial. BMJ Open. 2019;9(12). [doi: 10.1136/bmjopen-2019-032844]

20. Ghanim A, Silva MJ, Elfrink MEC, et al. Molar incisor hypomineralisation (MIH) training manual. Eur Arch Paediatr Dent. 2017;18(4):225-242. [doi: 10.1007/s40368-017-0293-9] [Medline: 28721667]

21. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building a software platform partnership. J Biomed Inform. 2019;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]

22. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. Jun 1999;8(2):135-160. [doi: 10.1177/096228029900800204] [Medline: 10501650]

23. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. Jun 2016;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]

24. Stata Statistical Software: Release 18. StataCorp. 2023. URL: https://www.stata.com/ [Accessed 2023-12-11]

25. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: https://www.R-project.org/ [Accessed 2025-11-25]

26. Porumb I, Leucuta DC, Nigoghossian M, et al. Caries lesion assessment using 3D virtual models. Medicina (Kaunas). 2023;59(12):2157. [doi: 10.3390/medicina59122157] [Medline: 37241175]

27. Boye U, Pretty IA, Tickle M, Walsh T. Comparison of caries detection methods using intra-oral digital photographs and visual examination. BMC Oral Health. 2013;13(1):6. [doi: 10.1186/1472-6831-13-6] [Medline: 23312001]

28. Boye U, Willasey A, Walsh T, Tickle M, Pretty IA. Comparison of photographic vs visual caries assessments. Community Dent Oral Epidemiol. 2013;41(6):526-533. [doi: 10.1111/cdoe.12049]

29. Mehta P, Byun R, Blinkhorn AS. Use of intra-oral camera for caries research. J Dent Health Oral Disord Ther. 2017;8. [doi: 10.15406/jdhodt.2017.08.00302]

30. Macey R, Walsh T, Riley P, et al. Fluorescence devices for the detection of dental caries. Cochrane Database Syst Rev. Dec 8, 2020;12(12):CD013811. [doi: 10.1002/14651858.CD013811] [Medline: 33319353]

31. Mendes FM, Hissadomi M, Imparato JCP. Effects of drying time and the presence of plaque on the in vitro performance of laser fluorescence in occlusal caries of primary teeth. Caries Res. 2004;38(2):104-108. [doi: 10.1159/000075933] [Medline: 14767166]

32. Carvalho RN, Letieri ADS, Vieira TI, et al. Accuracy of visual and image-based ICDAS vs micro-CT. Braz Oral Res. 2018;32:e60. [doi: 10.1590/1807-3107bor-2018.vol32.0060]

33. Schulz-Weidner N, Gruber M, Schraml EM, et al. Intraoral scans to aid communication in pediatric dentistry. Dent J. 2024;12(1):15. [doi: 10.3390/dj12010015]

34. Michou S, Lambach MS, Ntovas P, et al. Automated caries detection in vivo using a 3D intraoral scanner. Sci Rep. Oct 28, 2021;11(1):21276. [doi: 10.1038/s41598-021-00259-w] [Medline: 34711853]

35. Jones B, Chen T, Michou S, et al. Diagnostic agreement between visual examination and an automated scanner system with fluorescence for detecting and classifying occlusal carious lesions in primary teeth. J Dent. Oct 2024;149:105279. [doi: 10.1016/j.jdent.2024.105279] [Medline: 39121599]

36. Jones B, Lambach M, Chen T, et al. Dental caries detection in children using intraoral scans and deep learning. J Dent. Sep 2025;160:105906. [doi: 10.1016/j.jdent.2025.105906] [Medline: 40527440]

Abbreviations

HREC: Human Research Ethics Committee
ICC: Intraclass correlation coefficient
ICDAS: International Caries Detection and Assessment System
IOS: intraoral scanner
MIS BAIR: Melbourne Infant Study: BCG for Allergy and Infection Reduction
OR: odds ratio
REDCap: Research Electronic Data Capture