Title

Calibration and reliability testing of a novel asynchronous photographic plaque scoring system in young children

Journal

Journal of Public Health Dentistry, 83(1)

ISSN

0022-4006

Authors

Avenetti, David M
Martin, Molly A
Gansky, Stuart A
et al.

Publication Date

2023-03-01

DOI

10.1111/jphd.12557

Peer reviewed

HHS Public Access

Author manuscript

J Public Health Dent. Author manuscript; available in PMC 2024 March 01.

Published in final edited form as:

J Public Health Dent. 2023 March ; 83(1): 108–115. doi:10.1111/jphd.12557.

Calibration and reliability testing of a novel asynchronous photographic plaque scoring system in young children

David M. Avenetti, DDS, MSD, MPH[1], **Molly A. Martin, MD, MAPP**[2,3], **Stuart A. Gansky, MS, DrPH**[4], **Francisco J. Ramos-Gomez, DDS, MS, MPH**[5], **Susan Hyde, DDS, MPH, PhD**[4], **Rebecca Van Horn, RDH, BA**[1], **Bonnie Jue, DDS**[4], **Genesis F. Rosales, MSW**[3], **Nancy F. Cheng, MS, MS**[4], **Caroline H. Shiboski, DDS, PhD, MPH**[6]

[1]Department of Pediatric Dentistry, UIC College of Dentistry, University of Illinois Chicago, Chicago, Illinois, USA

[2]Department of Pediatrics, UIC College of Medicine, University of Illinois Chicago, Chicago, Illinois, USA

[3]Institute for Health Research and Policy, University of Illinois at Chicago, Chicago, Illinois, USA

[4]Department Preventive and Restorative Dental Sciences, UCSF School of Dentistry, University of California, San Francisco, California, USA

[5]Division of Preventative and Restorative Sciences, Section of Pediatric Dentistry, UCLA School of Dentistry, University of California, Los Angeles, California, USA

[6]Department of Orofacial Sciences, UCSF School of Dentistry, University of California, San Francisco, California, USA

Abstract

Objectives: The Simplified Oral Hygiene Index for Maxillary Incisors (OHI-MIS) is a novel plaque scoring system adapted for young children. This study describes calibration training and testing used to establish the inter- and intra-rater reliability for OHI-MIS measured from clinical photographs.

Methods: Two raters from the Coordinated Oral Health Promotion Chicago (CO-OP) and one from the Behavioral EConomics for Oral health iNnovation (BEECON) randomized controlled trials (RCTs) underwent calibration with gold standard raters, followed by annual re-calibration. Raters from CO-OP also completed inter-rater reliability testing; all three raters completed intra-rater reliability testing rounds. Photographs were obtained from children aged 9–39 months.

Results: All three raters achieved greater than 0.77 Lin's Concordance Correlation (LCC) versus gold standard consensus during calibration. All three raters had LCC ≥0.83 at recalibration 1 year later. CO-OP trial raters scored 604 photos (151 sets of 4 photographs); mostly both raters were

Correspondence David M. Avenetti, University of Illinois Chicago, College of Dentistry, Department of Pediatric Dentistry and Division of Prevention and Public Health Sciences, Chicago, Illinois, USA. avenetti@uic.edu.

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

CONFLICT OF INTEREST STATEMENT
The authors have no conflicts of interest to disclose.

somewhat/very confident in their scoring (≥89%), describing the most photos as "clear" (90% and 81%). The CO-OP inter-rater LCC for total OHI-MIS score was 0.86, changing little when low quality or confidence photos were removed. All three raters demonstrated high intra-rater reliability(≥0.83).

Conclusions: The OHI-MIS plaque scoring system on photos had good reliability within and between trials following protocol training and calibration. OHI-MIS provides a novel asynchronous plaque scoring system for use in young children. Non-clinicians in field or clinical settings can obtain photographs, offering new opportunities for research and clinical care.

Keywords

oral health; oral hygiene; public health dentistry; pediatric dentistry; epidemiologic methods; children; dental plaque

INTRODUCTION

Dental plaque biofilm and plaque accumulation are risk factors for dental caries and periodontal disease.[1, 2] While there are some limitations to the positive predictive value between plaque and caries, heavy visible plaque is consistently agreed to place a patient at very high caries risk.[3, 4] In young children specifically, plaque accumulation on the primary incisors is strongly associated with future caries.[5, 6]

Mechanisms for measuring plaque quantity and quality are described in the literature,[7,8] including the plaque indices developed by Greene and Vermillion,[9] Quigley and Hein,[10] and Silness and Loe.[11] Most research with plaque indices focuses on plaque control products, such as dentifrices, mouth rinses, toothbrushes, and other products.[12–16] Only limited efforts have been published that describe how plaque can serve as an objective measure of oral cleanliness, reflecting behaviors such as tooth brushing.[17–21]

Currently, there are no plaque indices for use in children with early primary dentition.[7, 8] The Simplified Oral Hygiene Index (OHI-S) was developed for use in children and adolescents aged 10–19 years, as well as adults.[22] The index measures debris and calculus on the surfaces of six permanent teeth—maxillary incisor (#8 labial), mandibular incisor (#24 labial), and four first molars (#3 and 14 buccal, 19 and 30 lingual). These sites correspond to World Dental Federation (FDI) teeth numbers 51 buccal, 71 labial, 56 buccal, 66 buccal, 76 lingual, and 86 lingual. Given that young children with primary dentition lack permanent first molars, developing a plaque index for use in children's primary teeth offers novel utility to assess oral hygiene, a function of health behaviors such as tooth brushing.

To address this gap, we created the Simplified Oral Hygiene Index of Maxillary Incisors (OHI-MIS) using the Debris Index (DI) criterion of the OHI-S to visually assess plaque on the labial (facial/buccal) surfaces of the four primary maxillary incisors. Maxillary primary incisors are the most involved teeth in early childhood caries (ECC).[23] Research assistants apply plaque disclosing agent prior to photographing the maxillary incisors, and trained and calibrated clinicians asynchronously evaluate the photographs to score the DI. Disclosing agents provide more effective plaque visualization which aids evaluating oral hygiene and improves effective debridement.[24, 25] The OHI-MIS is used for children younger than 4 years in two studies funded by the National Institute for Dental and Craniofacial Research (NIDCR) as part of the Oral Health Disparities in Children (OHDC) Consortium: the Coordinated Oral Health Promotion (CO-OP) Chicago and BEhavioral EConomics for Oral health iNnovation (BEECON) trials.

The purpose of this article is to describe the OHI-MIS calibration and reliability testing and describe the process for testing inter-rater reliability within and between OHDC Consortium trials. Specific objectives of this study are to describe: (a) the process for establishing the OHI-MIS reliability between the gold standard raters' consensus and study raters; (b) the assessment of photo quality and scoring confidence of photos used for plaque scoring; (c) the process and results for intra- and inter-rater reliability testing; and (d) variability in reliability by tooth location. Methods for evaluating inter- and intra-rater reliability within the CO-OP trial, where more than one clinician scores plaque, are described. Consideration is given to the effects of raters' confidence and assessment of the photo quality used for scoring plaque. The protocol describes the methodology of the OHI-MIS (Appendix A).

METHODS

CO-OP and BEECON trials: Study designs and populations

The OHDC Consortium's Coordinating Center (CC) at the University of California San Francisco (UCSF) conducted OHI-MIS reliability assessment for both CO-OP and BEECON trials. CO-OP, a Chicago-based two-arm cluster randomized controlled trial (RCT) testing efficacy of a family-focused community health worker oral health intervention to improve tooth brushing behaviors, measured OHI-MIS as an outcome for children aged 6–36 months.[26] BEECON, a Los Angeles-based RCT assessing monetary rewards versus delayed reward control through 6 months on early childhood caries preventive behaviors among children aged 6 months to <48 months from underserved (predominantly Latino) communities.[27] These studies were approved under the following IRBs: UCSF IRB #16–19968 (NCT03862443); UCSF IRB #17–23786 (NCT03576326); UCSF IRB #16–19920; UIC IRB #2015–0815; UIC IRB #2017–1090 (NCT03397589).

OHI-MIS overview

Simplified Oral Hygiene Index for Maxillary Incisors is a modification of the simplified oral hygiene index (OHI-S)[6] to assess children's oral hygiene. OHI-MIS is a mean score of the DI of the original visual OHI-S only on the buccal (facial/labial) surfaces of four deciduous maxillary incisors (MI). OHI-MIS is the primary outcome measure for the CO-OP trial and a secondary outcome for the BEECON trial. Research assistants (RAs) applied disclosing agents to MIs in participant homes or clinic. After application, RAs photographed disclosed teeth using digital cameras (iPhone 7) and uploaded photographs to a secure Research Electronic Data Capture (REDCap) server. Raters later reviewed photographs for quality assurance and scored them using OHI-MIS criteria.

Debris Index scoring criteria for the OHI-MIS are

0 = No debris or stain present.

1 = Soft debris covering not more than one-third of the tooth surface, or the presence of extrinsic stains without other debris regardless of surface area covered.

2 = Soft debris covering more than one-third, but not more than two-thirds, of the exposed tooth surface.

3 = Soft debris covering more than two-thirds of the exposed tooth surface.

Rater training and calibration

Two CC gold standard clinician raters (CHS and SH) held a 90-minute webinar for BEECON and CO-OP trial raters to review the OHI-MIS protocol and scoring criteria. Photographs in the webinar represented various clinical situations (i.e., missing teeth, partially erupted teeth, dental anomalies, and disclosing quality variations) that raters might encounter.

Thirty photographs from the BEECON pilot trial were used for calibration.[28] Data captured in the REDCap clinical trials management system[29] included the OHI-MIS scoring (0–3) for up to four surfaces, photo quality ("clear," "blurry," or "unreadable"), and the rater's confidence level. Two gold standard raters independently scored all photos, discussed discrepancies, and repeated scoring until achieving concordance of all scores denoted as the gold standard consensus reference.

Two CO-OP ("A" and "C") and one BEECON ("B") clinician raters independently scored the 30 photographs. OHI-MIS scores were calculated as the mean DI for up to four surfaces on each photograph. LCC was computed between the gold-standard consensus and each rater's scores.[30] LCC ≥0.75 was determined a priori as adequate calibration. Calibration batches to review photographs, answer questions, and refine and reinforce scoring criteria were to be repeated until adequacy was achieved. This process was achieved using 25 new photos randomly selected from the BEECON pilot trial with five photos randomly selected from the original 30. Raters were unaware the second set had repeats; all photos were scored independently.

Continuing inter-rater reliability—CO-OP Trial

Additional inter-rater reliability testing was conducted between CO-OP's two raters (A and C) to ensure concordance and readiness for individual coding. Initially, both raters scored batches of 15 randomly selected photographs collected during CO-OP trial enrollment, entering OHI-MIS scores separately. After scoring each batch, raters met to adjudicate scores when differences occurred. This involved reviewing photograph and score comparisons; differing scores were discussed until final score agreement was reached. LCC agreement between the raters' original scores was calculated. This was repeated at two-week intervals until LCC >0.75 was established. Additional inter-rater reliability testing was performed as above every 4–6 months throughout the trial to assess inter-rater consistency and agreement. In total, nine inter-rater reliability batches were completed with each round involving photos from 15–20 children.

Intra-rater reliability—CO-OP and BEECON trials

Intra-rater reliability testing was also conducted for the two CO-OP trial and one BEECON trial raters, which compared re-scoring photographs previously scored to evaluate each rater's internal consistency. Randomly selected photographs were re-assigned to raters and scored separately at least 4 weeks after initial scoring. Twenty to thirty photographs were selected for each intra-rater reliability batch. LCC was calculated between rescored and previously scored photos with LCC ≥0.75 again as the threshold. A second round of intra-rater reliability was conducted 12–18 months later.

Quality assurance for photograph collection

Recognizing that the ability to score photos accurately relies on proper disclosing technique and photo quality, raters provided regular monitoring and feedback about photograph quality to RAs photographing MIs. In the CO-OP trial, both raters screened all participant photographs within 3 days of collection. Feedback-targeted lighting, photo focus, photo vantage point, participant position, and quantity and quality of plaque disclosing. Quality assurance resulted in consistently high-quality photos. BEECON used its pilot trial to refine the image capture (photography) protocol, training, and technique. During the RCT, RAs and the project manager-reviewed photographs within 1 day to ensure clarity.

Analytic methods

Simplified Oral Hygiene Index for Maxillary Incisors score reliability was assessed with LCC and 95% confidence intervals (CIs) using a CC-developed custom macro in the SAS Statistical Analysis Software system (version 9.4, SAS Institute, Cary, NC); based on prior experience and evaluations with BEECON pilot trial data, a threshold of 0.75 was established a priori as demonstrating reliability. Supplemental reliability of tooth surface DI scores was assessed with kappa and weighted kappa statistics (with 95% CIs) despite kappa's well-known limitations such as agreement and skewed marginal distributions influencing estimates. The goal of calculating kappa statistics was to explore concordance at the individual tooth level, in addition to LCC for overall OHI-MIS score.

RESULTS

Baseline calibration with gold standard raters required two rounds of training and scoring to establish threshold LCC in May/June 2018. Following the initial scoring of 30 photos, including the two CO-OP and one BEECON raters, 29 photos with gold standard mean plaque score of 2.0 (SD = 0.52) were "scorable" with means ranging from 1.7 (SD = 0.68) to 2.2 (SD = 0.61). LCC ranged from 0.51 to 0.73, indicating no raters met the required threshold. Following additional training, the three raters scored a second batch of 30 photos with a gold standard consensus mean OHI-MIS score of 1.8 where the raters had mean scores ranging 1.9–2.0. In the second batch of photos, all three raters met the threshold of 0.75 with LCCs 0.77–0.89. Calibration versus gold standard consensus 1 year later included scoring 30 photos with a mean plaque score of 2.2 with the gold standard being 2.0. LCC ranged from 0.83 to 0.89, all exceeding the minimum 0.75 threshold requirement. While there was a slight tendency for raters to score higher than the gold standard raters' consensus, this difference was not statistically significant (Table 1 and Figure 1).

In the CO-OP trial, 151 participants were scored over 9 batches, yielding up to 604 scored photos. Among these, Rater C scored six teeth as missing or unreadable (1%), 55 as blurry (9.1%), 542 as clear (89.7%), and 1 data point was missing (<0.1%). Of the 151 children, Rater A indicated there was at least one clear photo for 133 children, while Rater C indicated 138 children had at least one clear photo. Rater A scored 12 teeth as missing or unreadable (2.0%), 100 as blurry (16.6%), 492 as clear (81.5%), and no data as missing. Despite some differences, there was general agreement in the rating of photo quality. Rater C was “very” (69.5%) or “somewhat confident” (22.4%) in more than 92% of ratings, while Rater A was “very” (56.6%) or “somewhat confident” (32.8%) in more than 89% of ratings. Though some photos were deemed “blurry,” teeth could still be scored from these photos. Photos that could not be scored were scored as “unreadable.” (Table 2).

Mean LCC between both CO-OP trial raters for total OHI-MIS score was 0.86 ranging 0.71–0.91 over nine batches of photo review. When either cases with missing data or cases with low quality photos were excluded, mean LCC remained at 0.86. When cases with low confidence were excluded, mean LCC was 0.88. Only batch 1 (LCC = 0.73) and batch 5 (LCC = 0.71) fell marginally below the threshold of 0.75. Excluding low quality and/or low confidence showed minimal effect on the overall LCC, indicating that exclusion was not necessary for analysis.

In the CO-OP trial, mean Simple Kappa Coefficient was 0.66 (range: 0.41–0.76) and the mean Weighted Kappa Coefficient was 0.73 (range: 0.54–0.85) between two raters over nine batches of photos reviewed. Raters completed three rounds of photo review and adjudication before they were allowed to score independently. When either cases with missing data, cases with low quality, or low confidence were removed, the mean Kappa and Weighted Kappa Coefficients remained unchanged. When cases with low confidence were excluded, the mean Kappa Coefficient was 0.69 (range: 0.41–0.87) and the mean Weighted Kappa Coefficient was 0.75 (range: 0.52–0.90). Excluding low quality and/or low confidence showed minimal effect on the overall Kappa Coefficients, indicating that exclusion was not necessary for analysis (Figure 2).

In the CO-OP trial, Raters A and C were in agreement with the adjudicated score 84.5% and 91.1% of the time, respectively. When there were differences, there was no tendency for over- or underscoring by a particular rater. With the exception of two out of 572 scores for Rater C, the difference between the raters’ scores and adjudicated scores was only one level of plaque score. Pearson correlations were calculated to determine if error was more prevalent when scoring particular teeth (central incisors vs. lateral incisors), however correlation coefficients of >0.90 were achieved for all teeth. Although not statistically significant, greater variance was noted among scores for the central incisors for Rater C and lateral incisors for Rater A (Table 3).

Raters maintained a high level of intra-rater reliability. Raters C and A reviewed photos from the CO-OP trial. Rater C had an LCC of 0.83 and 0.95, while Rater A had LCC of 0.88 and 0.99 for the 2 batches. Using photos from the main BEECON trial, Rater B had an intra-rater LCC ≥0.96 for the batches reviewed (Table 4).

DISCUSSION

Study results suggest that the OHI-MIS can be used with good reliability to assess plaque in young children. The process includes capturing photos in the field or clinical setting, gold standard clinician rater calibration, and clinician inter-rater and intra-rater reliability verification and monitoring. The OHI-MIS protocol can be applied to sites with a single clinician rater as well as sites with more than one clinician rater. In the CO-OP trial, when raters were not concordant, the difference in plaque score was usually one category, and there was no directional error by tooth being scored (central vs. lateral incisor).

The OHI-MIS plaque index provides a novel method for scoring plaque in young children as young as six months of age. The OHI-MIS index would likely have the most utility for research purposes rather than routine clinical care; however, plaque disclosing can serve as a good opportunity for patient education. OHI-MIS has the added benefit of field setting implementation as well as clinical settings. Assessing plaque scores in the clinic environment can be challenging, particularly with young children due to limited cooperation from fear and limited attention span. Non-clinicians (e.g., research assistants) can be trained to capture quality photos in non-clinical environments, such as the person's home. Clinicians can evaluate saved photographs, scoring plaque later in batches. Multiple clinicians can also review photographs one or more times, without extra participant burden. Non-clinicians capturing photos in environments comfortable for children offers a new method for assessing plaque costing less than clinical assessment and ultimately providing similar results. There were no participants in either study for whom the plaque disclosing was unsuccessful.

Through OHI-MIS development and reliability testing, some challenges with scoring related to unique dental considerations, such as determining adequate tooth eruption for scoring, dental anomalies (such as fused or congenitally missing teeth), and cavitated lesions or restorations emerged. Some challenges were also seen with scoring related to photo quality, including positioning, lighting, blurriness, and adequate retraction, so regular screening of photo quality with timely feedback is needed to ensure staff collect adequate quality photos. Not surprisingly, participant cooperation posed a frequent challenge, particularly with the young age group included in these studies. Standardized approaches to scoring are described in the study protocol to address the clinical situations in a consistent manner. Photographs allow for these unique clinical situations to be discussed and verified, as needed, through asynchronous photograph review rather than relying on accurate real-time assessment with the participant.

Although this study showed good reliability, consideration must be given to both the surface area and extent of plaque accumulation. For example, in some cases, teeth stained with a uniform light pink covering the tooth surface due to early biofilm. In these situations, there was the potential for the participant to receive a high plaque score despite a relatively small thickness of plaque, while some participants had significant plaque accumulation with a small surface area. This, as well as other unique considerations, are further discussed in the OHI-MIS protocol (Appendix A). However, further validation studies are needed to explore the relationship between plaque surface area and plaque thickness in the context of the OHI-MIS. Disclosed plaque varied from early plaque to a thick debris, reflecting longstanding plaque accumulation. A general limitation of using plaque disclosing and measurement to assess oral health is that plaque may not be stable over time. The methods used in this study provide a momentary assessment of the child's plaque condition. Additional longitudinal studies exploring the relationship between plaque score and other proximal and distal oral health outcomes are needed.

Several quantitative methods are available to assess calibration. In this study, these include concordance of the total OHI-MIS as well as individual tooth score agreement. LCC evaluates total OHI-MIS values' closeness to the exact gold standard consensus values, which does not consider variability in scores among teeth within an individual, but is the outcome measure these two trials planned a priori to use. To further evaluate tooth level agreement, weighted Kappa scores, and tooth DI score concordance were evaluated. Our findings suggested tooth level variability was modest, so OHI-MIS provided stable inter- and intra-rater reliability evaluation. Reliability with scoring photograph quality and confidence was not an aim of this study, although data suggest photo quality and raters' confidence can be consistently high with proper quality assurance and training.

Clinicians and researchers can use these methods to evaluate plaque index as a proxy for oral health behaviors, such as brushing quality. This provides an objective measure of health behavior and means of research assessment. Plaque index is an intermediate variable between brushing behaviors and future caries which can be evaluated and intervened upon as primary caries prevention. Future studies with longitudinal data can explore plaque's predictive value on subsequent caries using the novel OHI-MIS instrument as well as describe plaque score variability and stability over time.

These findings provide a mechanism to efficiently score dental plaque in young children outside of the clinical setting in the context of several research studies. The findings also suggest potential pragmatic application in the clinical field through teledentistry. Given geographic and other barriers families may face accessing care, the ability to capture photos that clinicians subsequently review may provide pragmatic screening for high-risk oral health behaviors or conditions. Although research assistants disclosed plaque and captured photographs in these trials, caregivers may be able to photograph their children's MIs and transmit them for asynchronous review. This concept came to be reality during the COVID-19 pandemic in which many children were unable to access dental care due to clinic closures.

The primary conclusions drawn from this study are that plaque scoring with good reliability is feasible in young children; clinician raters can be properly trained and calibrated to score plaque consistently within and between study sites; and utilizing photographs for children at a young age, whether screening for caries or assessing plaque, may have applications for primary, secondary, and tertiary prevention as well as epidemiologic assessment. Furthermore, the ability for non-clinicians to capture photos could provide a pragmatic and cost-effective means of conducting research and executing oral health screening and promotion strategies such as interventions aimed at increasing tooth brushing and plaque control.

Author Manuscript
Author Manuscript
Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors wish to acknowledge Anna Sandoval and Nattanit “Eunice” Songthangtham from the University of Illinois Chicago and Sarit Helman and Helen Lindau from the University of California, San Francisco for their support in this research.

FUNDING INFORMATION

BEECON: US National Institutes of Health/National Institute of Dental and Craniofacial Research, Award Nos. UH2/UH3-DE025514 and Coordinating Center Grant No. U01-DE025507. CO-OP: US National Institutes of Health/National Institute of Dental and Craniofacial Research, Award Nos. UH2/UH3-DE025483.

Funding information

National Institute of Dental and Craniofacial Research, Grant/Award Numbers: U01-DE025507, UH2-DE025483, UH2-DE025514, UH3-DE025483, UH3-DE025514; University of Illinois; University of California, San Francisco

REFERENCES

1. de Carvalho FG, Silva DS, Hebling J, Spolidorio LC, Spolidorio DM. Presence of mutans streptococci and Candida spp. in dental plaque/dentine of carious teeth and early childhood caries. Arch Oral Biol. 2006;51(11):1024–8. [PubMed: 16890907]
2. Dos Santos MN, Dos Santos LM, Francisco SB, Cury JA. Relationship among dental plaque composition, daily sugar exposure and caries in the primary dentition. Caries Res. 2002;36(5): 347–52. [PubMed: 12399695]
3. Featherstone JD, Chaffee BW. The evidence for caries management by risk assessment (CAMBRA®). Adv Dent Res. 2018 Feb; 29(1):9–14. [PubMed: 29355423]
4. Chaffee BW, Featherstone JDB, Zhan L. Pediatric caries risk assessment as a predictor of caries outcomes: a retrospective cohort study. Pediatr Dent. 2017;39(3):219–32. [PubMed: 28583247]
5. Alaluusua S, Renkonen OV. Streptococcus mutans establishment and dental caries experience in children from 2 to 4 years old. Eur J Oral Sci. 1983;91(6):453–7.
6. Alaluusua S, Malmivirta R. Early plaque accumulation—a sign for caries risk in young children. Community Dent Oral Epidemiol. 1994;22(5PT1):273–6. [PubMed: 7813174]
7. Wei SHY, Lang NP. Periodontal epidemiological indices for children and adolescents: II. Evaluation of oral hygiene III. Clinical applications. Pediatr Dent. 1982;4(1):64–73. [PubMed: 6960328]
8. Pretty IA, Edgar WM, Smith PW, Higham SM. Quantification of dental plaque in the research environment. J Dent. 2005;33(3):193–207. [PubMed: 15725520]
9. Greene JC, Vermillion JR. Oral hygiene index: a method for classifying oral hygiene status. The simplified oral hygiene index. J Am Dent Assoc. 1960;61(2):172–9.
10. Quigley GA, Hein JW. Comparative cleansing efficiency of manual and power brushing. J Am Dent Assoc. 1962;65(1):26–9. [PubMed: 14489483]
11. Silness J, Löe H. Periodontal disease in pregnancy II. Correlation between oral hygiene and periodontal condition. Acta Odontol Scand. 1964;22(1):121–35. [PubMed: 14158464]
12 <j/>(a). Deery C, Heanue M, Deacon S, Robinson PG, Walmsley AD, Worthington H, et al. The effectiveness of manual versus powered toothbrushes for dental health: a systematic review. J Dent. 2004; 32(3):197–211. [PubMed: 15001285] (b)Yaacob M, Worthington HV, Deacon SA, et al. Powered versus manual toothbrushing for oral health. Cochrane Database Syst Rev. 2014;6:CD002281.
13. Riley P, Lamont T. Triclosan/copolymer containing toothpastes for oral health. Cochrane Database Syst Rev. 2013;12: Art. No. CD010514.

14. Van Strydonk DA, Slot DE, Van der Velden U, et al. Effect of a chlorhexidine mouthrinse on plaque, gingival inflammation and staining in gingivitis patients: a systematic review. J Clin Periodontol. 2012;39(11):1042–55. [PubMed: 22957711]
15. Van Leeuwen MP, Slot DE, Van der Weijden GA. The effect of an essential-oils mouth rinse as compared to a vehicle solution on plaque and gingival inflammation: a systematic review and meta-analysis. Int J Dent Hyg. 2014;12(3):160–7. [PubMed: 24720368]
16. Zhou N, Wong HM, Wen YF, McGrath C. Oral health status of children and adolescents with intellectual disabilities: a systematic review and meta-analysis. Develop Med Child Neurol. 2017;59: 1019–26. [PubMed: 28627071]
17. Broadbent JM, Thomson WM, Boyens JV, Poulton R. Dental plaque and oral health during the first 32 years of life. J Am Dental Assoc. 2011;142(4):415–26.
18. Murthy PS, Shaik N, Deshmukh S, Girish MS. Effectiveness of plaque control with novel pediatric oral hygiene need station (modified oral irrigation device) as compared with manual brushing and flossing: randomized controlled pilot trial. Contemp Clin Dent. 2018;9:170–3. [PubMed: 29875555]
19. Scheerman JFM, van Meijel B, van Empelen P, Kramer GJC, Verrips GHW, Pakpour AH, et al. Study protocol of a randomized controlled trial to test the effect of a smartphone application on oral-health behavior and oral hygiene in adolescents with fixed orthodontic appliances. BMC Oral Health. 2018;18(1):19. 10.1186/s12903-018-0475-9 [PubMed: 29415697]
20. Poklepovic T, Worthington HV, Johnson TM, Sambunjak D, Imai P, Clarkson JE, et al. Interdental brushing for the prevention and control of periodontal diseases and dental caries in adults. Cochrane Database Syst Rev. 2013;12: Art. No. CD009857.
21. Alm A On dental caries and caries-related factors in children and teenagers. Swed Dent J Suppl. 2008;195:7–63.
22. Greene JG, Vermillion JR. The simplified oral hygiene index. J Am Dental Assoc. 1964;68(1):7–13.
23. Seow WK. Early childhood caries. Pediatr Clin North Am. 2018; 65(5):941–54. [PubMed: 30213355]
24. Fasoulas A, Pavlidou E, Petridis D, Mantzorou M, Seroglou K, Giaginis C. Detection of dental plaque with disclosing agents in the context of preventive oral hygiene training programs. Heliyon. 2019 Jul 10;5(7):e02064. 10.1016/j.heliyon.2019.e02064 [PubMed: 31334380]
25. Mensi M, Scotti E, Sordillo A, Agosti R, Calza S. Plaque disclosing agent as a guide for professional biofilm removal: a randomized controlled clinical trial. Int J Dent Hyg. 2020;18:285–94. 10.1111/idh.12442 [PubMed: 32348624]
26. Martin MA, Zimmerman LJ, Rosales GF, Lee HH, Songthangtham N, Pugach O, et al. Design and sample characteristics of COordinated Oral health promotion (CO-OP) Chicago: a cluster-randomized controlled trial. Contemp Clin Trials. 2020; 1(92):105919.
27. Ramos-Gomez F, White JS, Lindau HE, Lin TK, Finlayson TL, Liu JX, et al. Family monetary incentives as a value-based care model for Oral hygiene: rationale and design of the BEhavioral EConomics for Oral health iNnovation (BEECON) trial. J Public Health Dent. 2020;80:S17–S26. 10.1111/jphd.12406
28. White JS, Ramos-Gomez F, Liu JX, Jue B, Finlayson TL, Garza JR, et al. Monetary incentives for improving smartphone-measured oral hygiene behaviors in young children: a randomized pilot trial. PLoS One. 2020;15(7):e0236692. 10.1371/journal.pone.0236692 [PubMed: 32730310]
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81. [PubMed: 18929686]
30. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989;45(1):255–68. [PubMed: 2720055]
31. Greene JC, Vermillion JR. The simplified oral hygiene index. J Am Dental Assoc. 1964;68:25–31.
32. Wei SHY, Lang NP. Periodontal epidemiological indices for children and adolescents: II. Evaluation of oral hygiene. Pediatr Dent. 1981;4(1):73–96.

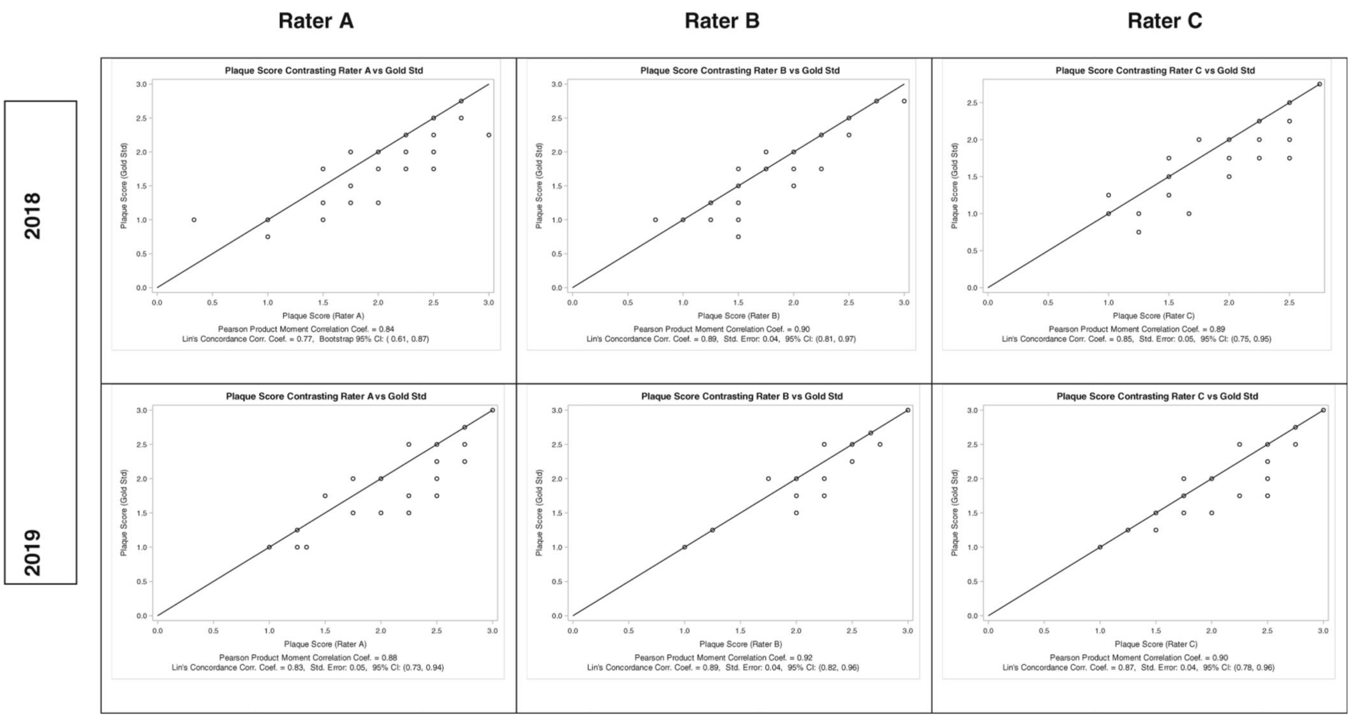

FIGURE 1.

Plaque calibration scatterplot and LCC results for each rater versus gold standard consensus.

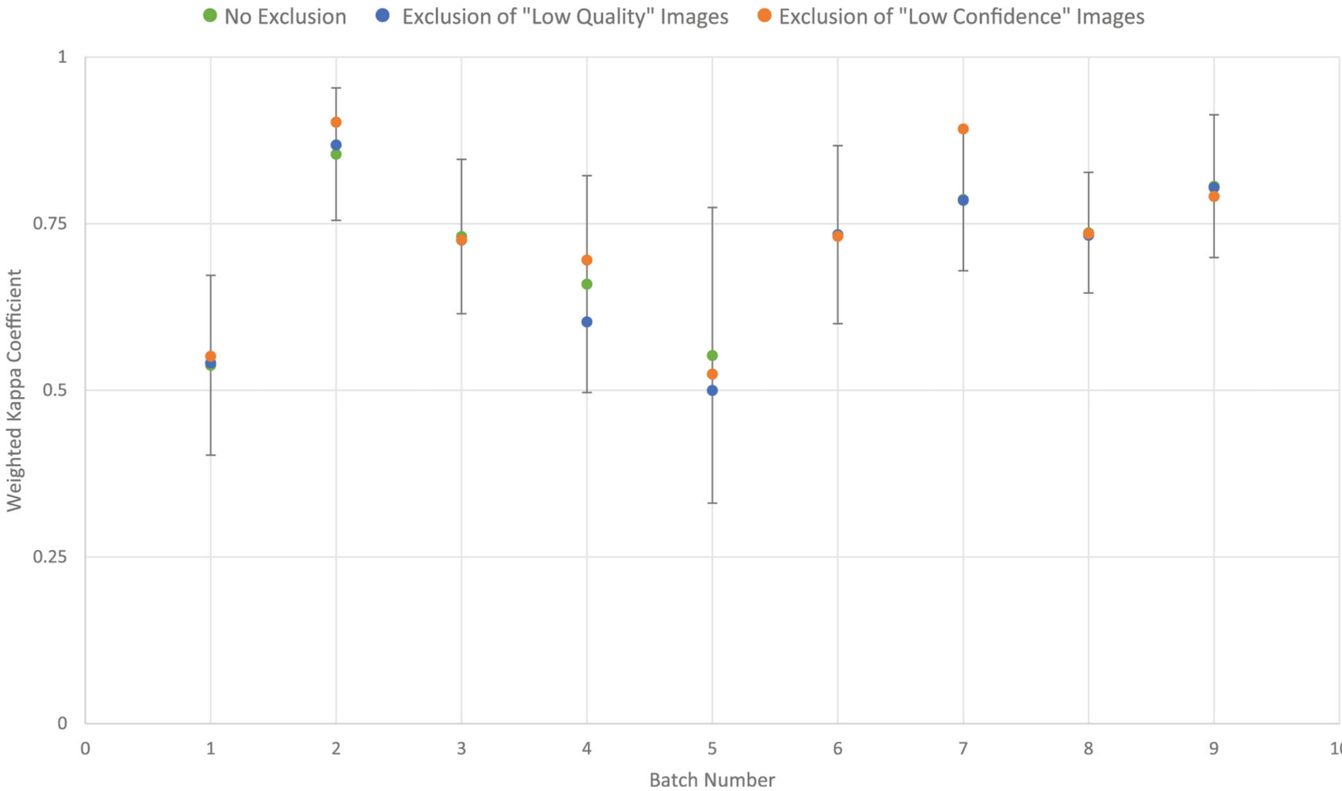

FIGURE 2.

Weighted Kappa coefficients between CO-OP trial raters without image exclusion and with exclusion of “low quality” and “low confidence” images (95% CI).

TABLE 1

Lin's concordance correlation between each rater and gold standard consensuss.

	April 2018		June 2018		June 2019	
	LCC	**Mean (SD)**	**LCC**	**Mean (SD)**	**LCC**	**Mean (SD)**
Gold Standard	Reference[a]	2.0 (0.52)	Reference[b]	1.8 (0.56)	Reference[a]	2.0 (0.58)
Rater A (CO-OP)	0.73[a]	1.9 (0.62)	0.77[a]	2.0 (0.63)	0.83[a]	2.2 (0.55)
Rater B (BEECON)	0.64[a]	2.2 (0.61)	0.89[b]	1.9 (0.54)	0.89[a]	2.2 (0.53)
Rater C (CO-OP)	0.51[a]	1.7 (0.68)	0.85[b]	2.0 (0.54)	0.87[a]	2.2 (0.55)

[a] $N = 29$ photos scored.

[b] $N = 30$ photos scored.

TABLE 2

CO-OP Trial Raters' evaluations of photo quality and confidence with scoring.

	Rater A	Rater C
No. of teeth (151 children; 9 batches)[a]	604	604
Photo quality		
No. of children with at least one tooth scored as clear	133	138
Missing teeth or Unreadable	2% (12)	1% (6)
Blurry	17% (100)	9% (55)
Clear	81% (492)	90% (542)
Missing data	0% (0)	<1% (1)
Confidence		
Very	57% (342)	70% (420)
Somewhat	33% (198)	22% (135)
Not very	5% (29)	2% (15)
Tooth missing	5% (32)	5% (30)
Missing data	1% (3)	<1% (4)

[a]Four photos per child.

TABLE 3

CO-OP Trial raters' scores relative to adjudicated scores.

	Rater A, *N* = 572	Rater C, *N* = 573
Rater's score relative to adjudicated score		
Underscored by 3	0% (0)	0% (0)
Underscored by 2	<1% (1)	0% (0)
Underscored by 1	3% (17)	9% (52)
Same score	91% (521)	84% (484)
Overscored by 1	6% (32)	6% (37)
Overscored by 2	<1% (1)	0% (0)
Overscored by 3	0% (0)	0% (0)
Pearson correlation between each rater's score and adjudicated score		
OHI-MIS	0.96	0.93
All Teeth	0.94	0.90
Tooth D (FDI tooth 12)	0.91	0.90
Tooth E (FDI tooth 11)	0.96	0.82
Tooth F (FDI tooth 21)	0.96	0.94
Tooth G (FDI tooth 22)	0.93	0.88

TABLE 4

Intra-rater OHI-MIS scoring reliability.[a]

	Batch 1	Batch 2	Overall
CO-OP Trial			
Dates	January 2019	January 2020	Both sets
No. of teeth	80	72	152
No. of children	20	18	38
Mean OHI-MIS[b]	2.1	2.0	2.1
Rater A	0.83[b]	0.95[c]	0.92[c]
Rater C	0.88[c]	0.99[c]	0.96[c]
BEECON Trial			
Dates	December 2018	June 2020	Both sets
No. of teeth	117	120	237
No. of children	30	30	60
Mean OHI-MIS[b]	2.0	2.2	2.1
Rater B	0.97	0.96	0.97

[a]LCC is for total OHI-MIS.

[b]Mean OHI-MIS from the first rating.

[c]LCC between the rater's score and adjudicated score.