University of Dundee

Reproducibility of Subtraction Radiography in Monitoring Changes in Approximal Carious Lesions in Children

Al-Sane, Mona; Ricketts, David N.; Mendes, Fausto M.; Altarakemah, Yacoub; Deery, Christopher; Innes, Nicola

Published in:
International Journal of Paediatric Dentistry

DOI:
[10.1111/ipd.12638](https://doi.org/10.1111/ipd.12638)

Publication date:
2020

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Al-Sane, M., Ricketts, D. N., Mendes, F. M., Altarakemah, Y., Deery, C., Innes, N., & Rollings, S. (2020). Reproducibility of Subtraction Radiography in Monitoring Changes in Approximal Carious Lesions in Children: An In Vivo Study. *International Journal of Paediatric Dentistry*, 30(5), 587-596. <https://doi.org/10.1111/ipd.12638>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Reproducibility of Subtraction Radiography in Monitoring Changes in Approximal Carious Lesions in Children: An In-Vivo Study

Short Title: Subtraction Radiography in Monitoring Caries

Mona Al-Sane^{1*}, David N. Ricketts², Fausto M. Mendes³, Yacoub Altarakemah⁴, Christopher Deery⁵, Nicola Innes⁶, Sam Rollings⁷

¹Kuwait University Faculty of Dentistry, Department of Developmental and Preventive Sciences, Kuwait City, KW.

²Dundee Dental Hospital and School, University of Dundee, Section of Operative Dentistry, Fixed Prosthodontics and Endodontology, Dundee, UK.

³School of Dentistry, University of Sao Paulo, Sao Paulo, Department of Pediatric Dentistry, Sao Paulo, BR.

⁴Kuwait University Faculty of Dentistry, Department of Restorative Sciences, Kuwait City, KW

⁵School of Clinical Dentistry, University of Sheffield, Unit of Oral health, Dentistry and Society, Sheffield, UK.

⁶ Dundee Dental Hospital and School, University of Dundee, Department of Paediatric Dentistry, Dundee, UK.

⁷Aberdeen Dental Hospital and Institute of Dentistry, Department of Restorative Dentistry, Aberdeen, UK.

*Corresponding Author

Mona Al-Sane BChD, MS

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/IPD.12638](#). This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

This article is protected by copyright. All rights reserved

Department of Developmental and Preventive Sciences, Faculty of Dentistry, Kuwait University, P.O.
Box: 24923, Safat, 13110, Kuwait City, Kuwait
Tel: +965 24986720
Fax: +96525326049
E-mail: mona.alsane@hsc.edu.kw

Author Contributions

MA designed the study, developed the data collection forms. Participated in data collection, wrote the manuscript. DR designed the study, developed the data collection forms, subtracted all radiographs included in the study, participated in data collection, helped with manuscript writing, reviewed and edited the manuscript. FM conducted all statistical analysis for this study, reviewed and edited the manuscript. YA converted radiographs included in the study into digital images, participated in the data collection, reviewed and edited the manuscript. CD, NI and SR participated in the data collection, reviewed and edited the manuscript.

Keywords: Caries, Proximal caries, Caries progression, Subtraction radiography, Reproducibility

Data Accessibility

Data available on request from the corresponding author.

Funding

This was an unfunded study

Word Count

4483 (excluding references and tables)

Conflict of Interest

The authors have no conflicts of interest to declare.

DR. MONA AL-SANE (Orcid ID : 0000-0003-4940-923X)

DR. FAUSTO MEDEIROS MENDES (Orcid ID : 0000-0003-1711-4103)

Article type : Original Article

Reproducibility of Subtraction Radiography in Monitoring Changes in Approximal Carious Lesions in Children: An In Vivo Study

Summary

Background: Laboratory studies have shown that digital subtraction radiography (DSR) can be a more effective tool, than conventional radiography, in monitoring changes in carious lesions. The clinical performance of the technique, however, has not been sufficiently researched.

Aim: To compare the reproducibility of DSR to that of bitewing radiographs, in monitoring changes in approximal caries in the mixed dentition. A secondary aim was to determine if assessment outcomes differed as a function of the method used.

Design: Six assessors evaluated 310 lesions first on bitewings, then with DSR. The overall reproducibility was evaluated via intra-class correlation coefficient (ICC). Intra and inter-rater reproducibility, were assessed using weighted Kappa. Paired t-test was used to assess differences in the reproducibility across methods.

Results: The overall reproducibility for DSR was (ICC=0.47, 95% CI=0.31-0.56). Intra and Inter-rater reproducibility were (0.65, 0.44) respectively. The overall reproducibility for bitewings was (ICC=0.45, 95% CI=0.42-0.57). Intra and Inter-rater reproducibility were

(0.71, 0.46) respectively. Differences in the reproducibility across methods were not statistically significant. Significantly more lesions were scored as progressed using DSR.

Conclusions: The reproducibility of DSR in monitoring changes in approximal caries, is comparable to that of bitewings. Additionally, DSR detected higher proportion of progression than bitewing assessments.

Keywords: Caries, Proximal caries, Caries progression, Subtraction radiography, Reproducibility

Introduction

Cariology has evolved dramatically over the past few decades, resulting in a paradigm shift in caries management. Over time, holistic minimally intervening strategies, with a more therapeutic approach to managing caries, have replaced traditional management philosophies focusing on restoring cavitated carious lesions. Currently, there is more emphasis on using caries risk assessment tools, preventive and therapeutic strategies, early lesion detection, and periodic surveillance. When intervention is required, more conservative, minimally invasive dentistry is advocated.¹⁻³

In contemporary dental practice, carious lesions are detected usually via a meticulous clinical visual examination of clean and dry teeth, supplemented with bitewing radiographs.⁴⁻¹⁰ However, on approximal surfaces, radiography can detect carious lesions at an early stage before they become clinically visible.¹¹⁻¹³ Preventive measures can arrest lesions at that early stage. Thus, lesion monitoring with sequential radiographs is needed to determine if lesions have progressed.¹⁴ Sequential radiographs are also useful for monitoring mineral content changes of lesions, when deeper, more established lesions are being managed with minimally invasive options, such as stepwise excavation and selective caries removal.¹⁵ Therefore, radiographic monitoring of carious lesion progression has become an integral component of practicing dentistry in the current era of minimally invasive dentistry.

Whilst viewing sequential bitewing radiographs side-by-side is a common approach for monitoring approximal lesions, subtle changes in the mineral content are often difficult to detect. Therefore, it is sometimes challenging to know with certainty whether a lesion has progressed or not. Studies have shown poor reproducibility for both inter- and intra-rater in radiographic diagnosis of carious lesions,¹⁶⁻¹⁸ which may result in different treatment decisions by different clinicians (or even by the same clinician) at different times. Several laboratory studies have suggested that digital subtraction radiography (DSR) can more accurately track the progression of approximal¹⁸ and occlusal^{19, 20} demineralisations than sequential side-by-side bitewing assessments. Studies have also shown that DSR has higher reproducibility in laboratory²⁰ and clinical studies.^{17, 21, 22}

DSR is an imaging technique that offers greater visualization of subtle changes in mineralised structures. In dentistry, the technique has been used for monitoring healing of maxillofacial fractures,²³ monitoring alveolar bone density changes in periodontal disease,^{24,25} diagnosing root fractures,²⁶ and monitoring changes in carious lesions.²⁰⁻²² The technique is easily applied to sequential digital radiographic images taken with similar X-ray projection

geometry.

A digital image consists of an array of pixels, each given a greyscale value. If two digital images taken at two different times are digitally superimposed, bespoke software can subtract the corresponding greyscale values. A progressed lesion with a net mineral loss will appear darker, one that has stayed the same will show no changes, and one with mineral gain will have a lighter (whiter) shade. A prerequisite for DSR use, however, is having sequential radiographs taken at identical X-ray projection geometries. This is difficult to achieve in everyday clinical practice, even with the use of individualised film holders. Variations in the bite-force can induce discrepancies in the X-ray projection geometry. Therefore, image correcting (warping) software have been introduced to DSR. These software help correct (warp) some level of variation in the X-ray projection geometry, which makes subtracting sequential non-standardised digital radiographs possible clinically.¹⁷

The literature has not sufficiently explored the clinical performance of DSR in detecting changes in carious lesions. There is only one published clinical study that evaluated the reproducibility of DSR in detecting changes in carious lesions in permanent teeth.¹⁷ No published studies have been conducted on primary teeth. Thus, the aim of the present study was to assess the reproducibility of DSR in detecting changes in approximal carious lesions in children in the mixed dentition and compare that to the reproducibility of bitewing radiographs, viewed side by side, in detecting such changes. A secondary aim of the study was, to determine if assessment outcomes differed as a function of the method used.

Material and Methods

This study was conducted using radiographs from the dental records of paediatric patients registered at the Kuwait University Dental Clinics (KUDC). Ethical clearance for this study was obtained from the Health Science Centre Ethical Clearance Committee, Kuwait University.

Selected Radiographs

The records of all patients admitted to the paediatric dental clinic over two consecutive years were screened for bitewing radiographs meeting the following criteria: patients in mixed dentition stage (6–12 years old), with sequential bitewing radiographs 6–18 months apart, of a good diagnostic quality, with no or minimal approximal overlap between most teeth. The screening process yielded 201 pairs of sequential bitewings. Of those, 122

pairs (60%), taken from 61 patients, were subtractable using DSR and were included in the current study. The mean patient's age included in the study was 8.3 years (SD = 2.1). The mean time difference between sequential radiographs was 7.1 months (SD = 1.6).

Once selected, all radiographs were coded and anonymised to all assessors involved in this study. Bitewings included in the study were taken with bitewing Rinn XCP film holders (Dentsply, Pennsylvania, USA), with no further standardisation. E-speed films were used (Eastman Kodak, Rochester, New Jersey, USA) in size 0 or size 1. Planmeca, Prostyle Intra (Planmeca, Helsinki, Finland) was used for X-rays. It was set at 66 kV, 8 mA, and 0.2-s exposure time. An automatic developer processed all films.

Selected Surfaces

From the selected 122 radiograph pairs, 310 approximal tooth surfaces were included in this study, of which 172 were on primary teeth, and 138 were on permanent teeth. Some of the included surfaces were sound, while others had carious lesions of varying severity.

Inclusion criteria were: approximal surface unrestored at baseline and follow-up, and in contact with a natural tooth surface at baseline and follow-up.

Image Photography and Image Manipulation

The selected 122 bitewings pairs were conventional images. Digitizing those radiographs was necessary, to allow them to undergo subtraction using DSR software. Additionally, assessors involved in the current study were located in different parts of the world. It was therefore decided to send all images in digital format in Microsoft Power Point files.

Bitewing pairs were converted into digital images by one of the authors (YT) using digital photography at a 3456×2304 pixels resolution, BIT depth 24, and 72 dpi horizontal and vertical resolution. Images were saved in the 'JPEG' format. Digital images of the bitewing pairs were saved in a Microsoft PowerPoint presentation with the surfaces to be assessed identified on the PowerPoint slides.

Digital images of the bitewings were manipulated with Corel PaintShop Photo Pro X3 (Ottawa, Canada). Each image was converted into a 256-greyscale image, the 'blur more' function was applied to the image, which was resized to 860×570 pixels and saved as a 'BMP' file. Compare software (Dental Health Unit, University of Manchester, UK), which runs as a plug-in to the Image Tool version 1.23 software (University of San Antonio, Texas, USA),²⁷ was used to create subtraction images by subtracting the baseline bitewings from the

corresponding follow-up radiograph, as described by Ricketts et al. (2007).²⁰ All subtractions were carried out by one of the authors (DR), who was experienced in using the Image Tool and Compare software. The resulting images were placed in separate PowerPoint presentations. Slides in both PowerPoint presentations (the side-by-side bitewings and the DSR images) were randomly coded using a random number table²⁸ and then sent to the assessors for scoring.

Reproducibility Assessment

Six assessors (three paediatric dentists and three prosthodontists/ restorative dentists) received the PowerPoint presentations with an information pack describing the DSR technique and explaining the process of interpreting subtraction images. Additionally, information on the aim of the study, and what was expected from assessors, were also included. Assessors scored the side-by-side digital bitewings and the DSR images on two separate occasions. For each surface, lesion progression was recorded using a five-point certitude scale, with 0 = Definitely no progression; 1 = Probably no progression; 2 = Unsure; 3 = Probably progressed; 4 = Definitely progressed.²⁰ Fifty percent of the surfaces assessed at baseline were randomly selected and re-examined two weeks later to determine the intra-rater reproducibility.

Statistical Analysis

For most of the analyses, the statistical unit was the score given to each lesion by the six different assessors, using two different methods (side-by-side bitewings and DSR image). Each score was clustered within each dental surface, and each dental surface was clustered within the different assessors, because dental surfaces were examined several times by different assessors. Thus, statistical methods appropriate to the clustered nature of these data were used. In performing the analyses, the following three hierarchical levels were considered: first level, type of assessment (DSR vs. bitewings); second level, dental surfaces (mesial vs. distal); third level, assessors (paediatric dentists vs prosthodontists or restorative dentists).

First, a descriptive analysis of the data was conducted. The six assessors' scores were pooled by the method. A multilevel Poisson regression analysis was conducted to determine whether scores differed by method; all scores were considered, and therefore, the outcome was a count variable.

With regard to the detection of progression, the assessment outcomes were then

categorised into one of two categories: (1) surfaces diagnosed with progression (combining both probable and definite), and (2) surfaces diagnosed with no progression (combining both probable and definite). Surfaces classified as ‘unsure’ were treated as missing values.

Thereafter, univariate multilevel Poisson regression analyses were conducted to evaluate the association between the assessment outcome and variables that could influence that outcome.

Explanatory variables included in the analysis were:

- Assessment method (DSR vs. bitewing)
- Dental surface (mesial vs. distal)
- Tooth type (primary vs. permanent)
- Baseline lesion depth (sound, enamel, or dentin)
- Specialty (paediatric dentists vs. prosthodontists or restorative dentists)

The variables that significantly predicted progression were entered into a univariate multiple regression model; prevalence ratio values (PR) and their respective 95% confidence intervals (CIs) were calculated.

Additionally, the difference between the methods was evaluated by multilevel Poisson regression in a stratified analysis that included all other variables that could influence the assessment outcome (dental surface examined, tooth type, baseline lesion depth, and observer’s speciality).

Reproducibility Assessment

Intra- and inter-rater reproducibility was assessed using raw scores and dichotomised scores. Overall reproducibility of the six assessors was evaluated via the intra-class correlation coefficient (ICC). Intra- and inter-rater reproducibility (2×2 assessors) were then assessed using weighted Kappa values and a quadratic approach (for raw scores) or unweighted Kappa values (for dichotomised scores). Paired t-tests were conducted to determine whether mean Kappa values differed between methods.

Interpretation of Reproducibility Indices

Overall reproducibility was assessed via ICC. As per Fleiss (1981),²⁹ ICC values of 0.40 or lower are poor, values between 0.40 and 0.59 are fair, values between 0.60 and 0.74 are good, and values above 0.75 are excellent. Intra- and inter-rater reproducibility were assessed via weighted Kappa values. As per Landis and Loch (1977),³⁰ Kappa values less than 0.2 are poor, values between 0.21 and 0.40 are fair, values between 0.41 and 0.60 are moderate, values between 0.61 and 0.80 indicate substantial agreement, and values above

0.81 indicate an almost perfect agreement.

Results

The findings indicate that lesion scoring of progression or arrest differed significantly across methods ($p < 0.001$) (Table 1). When using DSR, a higher percentage of lesions were scored as progressed (36%) compared to side-by-side bitewing assessment (31%).

Additionally, fewer lesions were scored as 'unsure' when DSR was used (12.5%) compared to when side-by-side bitewing assessment was used (14.5%).

Prediction of Progression

The outcome of the univariate multilevel Poisson regression analysis indicated that the bitewing method found significantly fewer cases of progression (36.6%) than the DSR method (41.4%) (Table 2). Additionally, the following variables were significantly associated with progression scores: type of surface, type of tooth, and baseline depth. Assessors scored a significantly higher number of lesions as progressed, when the surface was distal (44%), when the tooth was primary (50.4%), and when the baseline depth was in dentin (80.6%) or enamel (55.2%), as opposed to surfaces that were sound at baseline and developed a lesion at follow-up.

However, when controlling for the other variables, the method did not significantly predict a lesions' assessment outcome (progression vs. arrest) (Table 3). Only the type of tooth ($p < 0.001$) and baseline depth ($p < 0.001$) presented a significantly higher percentage of progression. Lesions on primary teeth, whether in enamel or dentin, were more likely to be scored as 'progressed' regardless of the assessment method used. However, sound surfaces at baseline that developed carious lesions at follow-up were less likely to be identified by either method.

The differences between the two assessment methods were evaluated by multilevel Poisson regression in a stratified analysis considering all other influencing variables (Table 4). The results indicate that assessors using the DSR method were significantly more likely to score carious lesions as progressed if the lesions were on permanent teeth ($p = 0.047$), when the assessed surface was sound at baseline ($p < 0.001$), and when surfaces affected by caries were mesial ($p = 0.017$). Thus, DSR was better at picking up progression on permanent teeth, newly developed carious lesions on surfaces that were sound at baseline, and when progression occurred on a mesial surface.

Reproducibility Assessment

Raw Lesion Progression Scores

- A. *DSR method.* Overall reproducibility for the scores rated using the DSR method was poor (ICC = 0.39, 95% CI [0.31–0.48]). Intra-rater reproducibility weighted Kappa values ranged from 0.56–0.79; the mean weighted Kappa value was 0.65 ($SD=0.1$). Therefore, intra-rater reproducibility was substantial. Inter-rater reproducibility weighted Kappa values ranged from 0.23–0.64; the mean weighted Kappa value was 0.44($SD=0.13$). Therefore, inter-rater reproducibility was moderate (Table 5).
- B. *Bitewing method.* Overall reproducibility for the scores rated using the bitewing method was fair (ICC = 0.49, 95% CI [0.42–0.57]). Intra-rater reproducibility weighted Kappa values ranged from 0.61–0.81; the mean weighted Kappa value was 0.70 ($SD=0.07$). Therefore, intra-rater reproducibility was substantial. Inter-rater reproducibility weighted Kappa values ranged from 0.19–0.63; the mean weighted Kappa value was 0.48($SD=0.15$). Therefore, inter-rater reproducibility was moderate (Table 5).

Comparison between DSR and Bitewing Reproducibility for Raw Lesion Scores

Intra-rater reproducibility while using the DSR method did not differ significantly from intra-rater reproducibility using the bitewing method ($p = 0.324$). Similarly, inter-rater reproducibility while using the DSR method did not differ significantly from inter-rater reproducibility while using the bitewing method ($p = 0.318$).

Dichotomised Lesion Progression Scores

- A. *DSR method.* Overall reproducibility for scores using the DSR method was fair (ICC = 0.47, 95% CI [0.39–0.56]). Intra-rater reproducibility weighted Kappa values ranged from 0.58–0.82; the mean weighted Kappa value was 0.65($DS=0.09$). Therefore, intra-rater reproducibility was substantial. Inter-rater reproducibility weighted Kappa values ranged from 0.15–0.61; the mean weighted Kappa value was 0.44. Therefore, inter-rater reproducibility was moderate (Table 6).
- B. *Bitewing method.* Overall reproducibility for the scores rated using the Bitewing method was fair (ICC = 0.54, 95% CI [0.47 to 0.61]). Intra-rater reproducibility weighted Kappa values ranged from 0.58–0.80; the mean weighted Kappa value was 0.71($SD=0.09$). Therefore, intra-rater reproducibility was substantial. Inter-rater

reproducibility weighted Kappa values ranged from 0.20–0.72; the mean weighted Kappa value was 0.46($SD=0.15$). Therefore, inter-rater reproducibility was moderate (Table 6).

Comparison between DSR and Bitewing Reproducibility for Dichotomised Lesion Scores

Differences in the intra and inter-rater reproducibility across methods were not statistically significant ($p>0.05$).

Discussion

To the authors' knowledge, this is the first *in vivo* study investigating the reproducibility of DSR for monitoring changes in approximal carious lesions in the mixed dentition. A secondary aim of the study was to assess whether detecting changes in approximal caries differed as a function of method used (DSR vs. bitewings). The study was designed to assess DSR's diagnostic potential in detecting changes in carious lesions in everyday clinical practice. Therefore, the present study was conducted on radiographs of paediatric patients from existing patient records taken using Rinn (XCP) film holders with no further standardization. No new radiographs were taken for the purpose of this study. The radiographs used were for high caries risk children, in the mixed dentition. Carious surfaces, both in primary and permanent teeth, were assessed in the current study.

Ideally, when assessing the potential of caries diagnostic tests, both accuracy and reproducibility should be evaluated. Well performed accuracy studies on DSR's effectiveness in detecting changes in carious lesions require an objective quantification of disease levels or a 'gold standard', which can only be achieved in laboratory studies, as it requires examining histological sections of carious lesions, both at baseline and follow-up. Therefore, *in vivo* diagnostic studies on the subject are limited to reproducibility assessments¹⁷ and possibly subjective evaluation of changes in carious lesions.

The main finding of this study is that when monitoring changes in approximal carious lesions in the mixed dentition, the overall reproducibility of DSR was comparable to that of bitewings; both methods had a fair overall reproducibility. Further, DSR detected a larger proportion of carious lesions progressing on approximal surfaces compared to conventional side-by-side bitewing assessments.

A closer look at the reproducibility results revealed a substantial intra-observer reproducibility for the DSR assessments.³⁰ Six assessors, with variable levels of experience in interpreting DSR images, participated in the present study. They received no training other

than an information pack before the start of this study. A few *in-vivo* studies have reported intra-observer reproducibility of DSR for assessing changes in carious lesions over time. Wenzel et al.¹⁷ reported an almost perfect reproducibility (>0.81). Assessors participating in that study received training before the start of the study, which could explain the higher reproducibility results compared to the current study.

Additionally, Martignon et al.^{21,22} reported an almost perfect reproducibility (>0.81) of DSR for assessing changes in carious lesions. In their studies all assessments were conducted by one observer who was experienced in interpreting DSR images. Based on the above, one can speculate that DSR's intra-observer reproducibility would improve with training and experience. This is worth investigating in future studies.

Results of the present study revealed that intra-observer reproducibility of the bitewings assessments was comparable to that of DSR. There was no statistically significant difference across the methods. This observation is different from that reported in previous similar studies, where DSR has always shown statistically superior intra-observer reproducibility results than bitewing radiographs.^{17,21,22} However, different studies should be cautiously compared. Study design differences and observer's background and experience level can influence the results of individual studies. The above-cited studies were conducted under controlled research conditions. Customised film holders were used in two of those studies. Assessors involved in the cited studies were either experienced or received detailed training before the start of the study. Additionally, while previously studies were conducted on permanent teeth, the present study was conducted on mixed dentition teeth, with caries in both primary as well as permanent teeth being assessed. This may have affected our results; as the rate of caries progression varies from primary to permanent teeth. Furthermore, the dynamic changes due to the exfoliation of primary teeth and the eruption of new permanent teeth may have also affected the results of our study.

Analysis of data revealed that the inter-rater reproducibility was moderate for the DSR and bitewing methods, with no statistically significant difference. These values, however, were lower than those reported previously by Wenzel et al.,¹⁷ who reported that the inter-rater reproducibility was substantial for both DSR and bitewing radiographs. Unlike assessors in their study, those in the present study were not calibrated before the start of the study, which was intentional to reflect what would likely happen with both techniques in actual clinical practice. Lower inter-rater reproducibility values is a common observation in radiographic caries diagnosis.^{16,17} Individual variations in interpreting subtle radiographic changes are difficult to eliminate in everyday clinical practice, unlike the controlled research

environment, where calibration is possible.

The present study was not an accuracy one, yet some observed trends are worth highlighting here. Firstly, a crude analysis of the data revealed that using the subtraction method, assessors were significantly more likely to diagnose lesions as ‘progressed’ compared to when sequential side-by-side bitewings were used. Additionally, assessors had significantly fewer uncertainties about their scores when the subtraction method was used. However, once adjusted for confounders, these associations were not significant. Secondly, using subtraction, assessors were significantly more likely to diagnose changes in carious lesions on permanent than on primary teeth. Additionally, they were more likely to identify the development of new enamel lesions on surfaces that were sound at baseline. However, since true validation of the state of disease was not possible in this study, these observations do not substantiate the conclusion that DSR is more accurate, more sensitive to diagnosing progression in permanent teeth, or more sensitive for identifying newly developed enamel lesions. Such observations are worth investigating in future well designed accuracy studies, where the true state of disease can be quantified, for instance, by using micro CT in laboratory studies.

Laboratory studies have shown that DSR is more accurate and reproducible in detecting changes in carious lesions than bitewing radiographs.^{18,20} However, DSR is not used in caries diagnoses in everyday clinical practice. This is mostly due to the belief that total standardization of the X-ray projection geometry is essential for successful subtraction of sequential radiographs. This can be achieved with the use of customized film holders, which is often used in research settings. It is, however, not a practical option to use in routine patient care settings. XCP film holders used in the present study allowed for some, but not total, standardization of the projection geometry. Additionally, growth-related changes, exfoliation, and eruption of teeth in the mixed dentition made subtraction more difficult. To overcome these discrepancies, software with features that can correct such discrepancies was used, as described earlier. Additionally, teeth included in this study were subtracted in isolation from the surrounding structures. This helped compensate for discrepancies in the projection geometry and made subtracting sequential non-standardised bitewings possible. Wenzel et al.¹⁷ reported successfully subtracting non-standardised bitewings taken using a Kwik-Bite film holder. That study highlighted that experience is needed to optimally subtract non-standardised images.¹⁷ Currently, no published studies have reported on the level of discrepancy in the X-ray projection geometry that can be compensated for with the use of image warping software, and the effect of that compensation on the accuracy and

reproducibility of diagnosing changes in approximal caries. This is worth investigating in future studies. Additionally, no published studies have investigated the effect on the accuracy and the reproducibility of the diagnostic decisions when subtracting carious teeth in isolation from the surrounding structures. This is worth further exploration in future studies.

In the present study, close to 60% of the radiographs that met our inclusion criteria were subtractable. Assessing the subtractability of images taken using XCP film holders without further standardisation was beyond the scope of this study. However, this is worth exploring in future studies.

A strength of the present study was that it was designed to assess the clinical performance of DSR under conditions resembling those of everyday clinical practice. The use of non-standardised bitewings and the absence of calibration were all intentional. An additional strength is that intra-rater reproducibility was assessed using 50% of the sample. Most caries reproducibility studies are conducted using 10%–25% of the sample for re-assessments.^{17,20-22} Furthermore, to reduce recall bias, all re-assessments conducted in the present study were done two weeks after baseline assessments, with images being shuffled using a randomization table.²⁸ These measures added strength and credibility to the intra-observer reproducibility outcome of the current study.

One of the study limitations, however, was the lack of in-depth training of assessors before the start of the study. Assessors were based at different centres, at different geographic locations. This made conducting an in depth, live, training session difficult. Therefore, for practical reasons, we opted to use an instruction manual with details on how to interpret subtraction images. The manual was emailed to all assessors before the start of the study.

All assessors were experienced clinicians who relied on bitewings as an indispensable caries diagnostic tool in their clinical practice. Most, however, were not as familiar with DSR. We speculate that the outcome of DSR reproducibility could have been enhanced if assessors were more experienced or more trained in using the technique. Findings of the current study should, therefore, be viewed within the context of this limitation. This study is worth repeating with assessors receiving an in depth DSR training beforehand.

DSR is currently not available for commercial use. The present study is one of the few studies on the clinical performance of this technique in caries diagnosis. More research is needed for a better understanding of DSR's and image warping software performance in actual clinical practice. Future research should focus on assessing the effectiveness, cost-effectiveness, and limitations of this technique.

Technology is evolving rapidly. Over the past decade, digital radiography has become

an integral part of everyday clinical practice. Currently, most contemporary digital radiography systems are equipped with image-enhancing software. If, after enough clinical research, DSR is shown to be an effective technique in monitoring caries, subtraction and image warping software could potentially be incorporated into image enhancing software, making it possible for the technique to be within reach to dentists in clinical practice.

Conclusion

In the present study, the overall reproducibility DSR was comparable to that of sequential bitewings (viewed side-by-side) for monitoring changes in approximal caries in the mixed dentition. Both methods had fair overall reproducibility. Additionally, the intra-observer reproducibility was substantial for both methods, without a statistically significant difference. The inter-observer reproducibility was moderate for both methods, with differences that were not statistically significant. Further, DSR detected a higher proportion of progression in approximal caries than conventional side-by-side bitewing assessments. Thus, the current study's outcome draws attention to DSR's potential for enhancing the detection of changes in approximal carious lesions in mixed dentition.

Why this Paper is Important to Paediatric Dentists

- Digital subtraction radiography has the potential of enhancing the detection of changes in approximal carious lesions in the primary and permanent teeth.
- Enhancing detection of changes in carious lesions is important in the era of minimal intervention dentistry where early lesion diagnosis, surveillance, and conservative management have become essential components in the practice of contemporary paediatric dentistry.

References

1. Featherstone JD, Doméjean S. Minimal intervention dentistry: part 1. From 'compulsive' restorative dentistry to rational therapeutic strategies. *Br Dent J* 2012;**213**:441-445.
2. Ismail AI, Tellez M, Pitts NB, et al. Caries management pathways preserve dental tissues and promote oral health. *Community Dent Oral Epidemiol* 2013;**41**:e12-e40.
3. Pitts NB, Zero DT, Marsh PD, et al. Dental caries. *Nat Rev Dis Primers* 2017;**3**:17030.
4. Martignon S, Pitts NB, et al. Caries Care practice guide: consensus on evidence into practice. *Br Dent J* 2019;**227**:353-362.
5. Pitts NB, Ismail AI, Martignon S, Ekstrand K, Douglas GVA, Longbottom C. ICCMS™ Guide for Practitioners and Educators. 2014. [URL accessed at: (www.Doi.org/10.5281/zenodo.853106) on (08. 01 .2020)].
6. Wenzel A. Radiographic display of carious lesions and cavitation in approximal surfaces: advantages and drawbacks of conventional and advanced modalities. *Acta Odontol Scand* 2014;**72**:251-264.
7. Gimenez T, Piovesan C, Braga MM, et al. Visual inspection for caries detection: a systematic review and meta-analysis. *J Dent Res* 2015;**94**:895-904.
8. Schwendicke F, Tzschope M, Paris S. Radiographic caries detection: a systematic review and meta-analysis. *J Dent* 2015;**43**:924-933.
9. Chen R. The international caries detection and assessment system is a visual diagnostic system that is highly reproducible and accurate for coronal carious lesions detection but only moderately reproducible and accurate for assessing lesion progression. *J Evid Based Dent Pract* 2019;**19**:91-94.
10. Pontes LRA, Novaes TF, Lara JS, et al. Impact of the radiographic method on treatment decisions related to dental caries in primary molars: a before-after study. *Clin Oral Investig* 2019. doi: 10.1007/s00784-019-02844-w.

11. Ekstrand KR, Luna LE, Promisiero L, et al. The Reliability and Accuracy of Two Methods for Proximal Caries Detection and Depth on Directly Visible Proximal Surfaces: An in vitro Study. *Caries Res* 2011;**45**:93-99.
12. Cortes A, Ekstrand KR, Gamboa LF, González L, Martignon S. Caries status in young Colombian children expressed by the ICCMS™ visual/radiographic combined caries staging system. *Acta Odontol Scand* 2017;**75**:12-20.
13. Subka S, Rodd H, Nugent Z, Deery C. In vivo validity of proximal caries detection in primary teeth, with histological validation. *Int J Paediatr Dent* 2019;**29**:429-438.
14. Cortes A, Ekstrand KR, Martignon S. Visual and radiographic merged-ICDAS caries progression pattern in 2-6 years old Colombian children: Two-year follow-up. *Int J Paediatr Dent* 2018 Nov 15. doi: 10.1111/ipd.12448. [Epub ahead of print]. PMID: 30431189. 1.
15. Ricketts D, Innes N, Schwendicke F. Selective removal of carious tissue. *Monogr Oral Sci* 2018;**27**:82-91.
16. Hintze H, Wenzel A. Clinically undetected dental caries assessed by bitewing screening in children with little caries experience. *Dentomaxillofac Radiol* 1994;**23**:19-23.
17. Wenzel A, Anthonisen PN, Juul MB. Reproducibility in the assessment of caries lesion behaviour: a comparison between conventional film and subtraction radiography. *Caries Res* 2000;**34**:214-218.
18. Ferreira RI, Haiter-Neto F, Tabchoury CP, de Paiva GA, Bóscolo F. Assessment of enamel demineralization using conventional, digital, and digitized radiography. *Braz Oral Res* 2006;**20**:114-119.
19. Eberhard J, Hartman B, Lenhard M, Mayer T, Kocher T, Eickholz P. Digital subtraction radiography for monitoring dental demineralization. An in vitro study. *Caries Res* 2000;**34**:219-224.

20. Ricketts DN, Ekstrand KR, Martignon S, Ellwood R, Alatsaris M, Nugent Z. Accuracy and reproducibility of conventional radiographic assessment and subtraction radiography in detecting demineralization in occlusal surfaces. *Caries Res* 2007;**41**:121-128.
21. Martignon S, Ekstrand KR, Ellwood R. Efficacy of sealing proximal early active lesions: an 18-month clinical study evaluated by conventional and subtraction radiography. *Caries Res* 2006;**40**:382-388.
22. Martignon S, Ekstrand KR, Gomez J, Lara JS, Cortes A. Infiltrating/sealing proximal caries lesions: a 3-year randomized clinical trial. *J Dent Res* 2012;**91**:288-292.
23. Queiroz CS, Sarmiento VA, de Azevedo RA, de Oliveira TF, Bastos LC. A comparative study of internal fixation and intermaxillary fixation on bone repair of mandibular fractures through radiographic subtraction. *J Craniomaxillofac Surg* 2014;**42**:e152-e156.
24. Zaman MU, Nakamoto T, Tanimoto K. A retrospective study of digital subtraction technique to detect sclerotic changes in alveolar bone on intraoral radiographs of bisphosphonate-treated patients. *Dentomaxillofac Radiol* 2013;**42**:20130242.
25. Matsuda Y, Terauchi T, Murahira K, et al. Exposure time reduction of secondary radiographs used in digital subtraction radiography in detecting intrabony change. *Oral Radiol* 2014;**30**:20-26.
26. Queiroz PM, Nascimento HA, da Paz TD, Anacleto FN, Freitas DQ. Accuracy of digital subtraction radiography in the detection of vertical root fractures. *J Endod* 2016;**42**:896-899.
27. Ellwood R, Horner K, Alexander S, Davies R. A digital subtraction radiography investigation of upper first molar proximal bone density changes in adolescents. *J Periodontal Res*. 1998 Apr;**33**:172-7.
28. Daniel WW. Biostatistics: A Foundation for Analysis in the Health Sciences. New York: John Wiley and Sons; 1974.
29. Fleiss J. Statistical Methods for Rates and Proportions. New York: John Wiley; 1981.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159-174.

Tables

Table 1. Descriptive analysis considering all scores of progressions obtained by the two different methods

	DSR [†]	Bitewing	p*
All individual scores			<0.001
Definitely no progression	755 (40.6)	816 (43.9)	
Probably no progression	193 (10.4)	180 (9.7)	
Unsure	231 (12.4)	272 (14.6)	
Probably progressed	261 (14.0)	299 (16.1)	
Definitely progressed	409 (22.0)	276 (14.8)	
Not assessed [‡]	11 (0.6)	17 (0.9)	
Total	1860 (100.0)	1860 (100.0)	

[†] Digital subtraction radiography

[‡] Data not included in the analysis

* Obtained through the Wald test performed after multilevel Poisson regression analysis

Table 2. Univariate multilevel Poisson regression analysis to evaluate the association of the dichotomized assessment outcome (progression vs. no progression) and other variables

Explanatory variables	No Progression N (%)	Progression N (%)	Unadjusted PR [‡] (95% CI)	P
Method				0.024
DSR [†]	948 (58.6)	670 (41.4)	1.00	
Bitewing	996 (63.4)	575 (36.6)	0.88 (0.79–0.98)	
Surface				0.001
Mesial	1296 (63.8)	735 (36.2)	1.00	
Distal	648 (56.0)	510 (44.0)	1.21 (1.09–1.36)	
Type of tooth				<0.001
Primary	845 (49.6)	858 (50.4)	1.00	
Permanent	1099 (74.0)	387 (26.0)	0.51 (0.46–0.59)	
Baseline depth				<0.001
Sound	1561 (76.4)	481 (23.6)	1.00	
Enamel	283 (44.8)	348 (55.2)	2.32 (2.02–2.67)	
Dentin	100 (19.4)	416 (80.6)	3.39 (2.97–3.87)	
Specialty				0.229
Paediatric dentists	906 (56.4)	700 (43.6)	1.00	
Prosthodontics	1038 (65.6)	545 (34.4)	0.83 (0.61–1.13)	

[†] Digital subtraction radiography

[‡] Prevalence ratio

Table 3. Multiple multilevel Poisson regression analysis to evaluate the association of the methods and other explanatory variables with the classification of dental surfaces considered as a dichotomous outcome (no progression vs. progression)

Explanatory variables		Adjusted PR [‡] (95% CI)	p
Method			0.080
Type of tooth	DSR [†]	1.00	
	Bitewing	0.91 (0.81–1.01)	
	Primary	1.00	<0.001
	Permanent	0.76 (0.67–0.88)	
Baseline depth			<0.001
	Sound	1.00	
	Enamel	2.24 (1.95–2.58)	
	Dentin	2.91 (2.51–3.38)	

[†] Digital subtraction radiography

[‡] Prevalence ratio

Table 4. Univariate multilevel Poisson regression analysis to investigate differences between two methods for classifying progression of caries lesions on dental surfaces stratified for other variables

Explanatory variables	Surfaces classified as progressed		PR ^{‡,§} (95% CI)	p
	N (%)			
	DSR [†]	Bitewing		
Type of tooth				
Primary	456 (52.1)	402 (48.6)	0.93 (0.81–1.06)	0.260
Permanent	214 (28.8)	173 (23.3)	0.82 (0.67–1.00)	0.047
Baseline depth				
Sound	286 (28.2)	195 (19.0)	0.67 (0.56–0.81)	<0.001
Enamel	164 (49.2)	184 (61.7)	1.25 (1.01–1.55)	0.035
Dentin	220 (80.9)	196 (80.3)	0.99 (0.82–1.20)	0.944
Surface				
Mesial	405 (39.3)	330 (33.0)	0.84 (0.72–0.97)	0.017
Distal	265 (45.1)	245 (42.9)	0.94 (0.79–1.12)	0.517
Specialty				
Paediatric dentists	380 (45.4)	320 (41.6)	0.90 (0.78–1.05)	0.172
Prosthodontics	290 (37.1)	255 (31.8)	0.86 (0.72–1.01)	0.071

[†] Digital subtraction radiography

[‡] Prevalence ratio

[§] The category of reference for these comparisons is the subtraction method

Table 5. Inter- and intra-rater reproducibility of the assessments made by different methods considering all scores of progression of caries lesions

	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6
	DSR [†] method [*]					
Ex 1	0.586	0.463	0.510	0.507	0.537	0.281
Ex 2		0.560	0.523	0.383	0.420	0.226
Ex 3			0.747	0.528	0.595	0.435
Ex 4				0.620	0.638	0.264
Ex 5					0.788	0.276
Ex 6						0.583
ICC [‡] (95% CI)	0.392 (0.306 to 0.475)					
	Bitewing [*]					
Ex 1	0.713	0.327	0.628	0.694	0.647	0.523
Ex 2		0.666	0.366	0.332	0.271	0.190
Ex 3			0.811	0.585	0.565	0.506
Ex 4				0.752	0.525	0.504
Ex 5					0.612	0.511
Ex 6						0.640
ICC [‡] (95% CI)	0.491 (0.423 to 0.570)					

[†] Digital subtraction radiography

[‡] Intraclass correlation coefficient

^{*} Weighted Kappa values. Shaded cells correspond to intra-rater reproducibility and unshaded cells are related to inter-rater reproducibility

Mean (standard deviation) of Kappa values:

Inter-rater reproducibility

DSR = 0.439 (0.128); Bitewing = 0.478 (0.148); p = 0.318

Intra-rater reproducibility

DSR = 0.647 (0.096); Bitewing = 0.699 (0.074); p = 0.32

Table 6. Inter- and intra-rater reproducibility of the assessments made by different methods considering the scores of progressions as a dichotomous variable (no progression vs. progressed)

	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6
DSR [†] method [*]						
Ex 1	0.596	0.593	0.522	0.498	0.609	0.212
Ex 2		0.602	0.524	0.443	0.572	0.146
Ex 3			0.708	0.471	0.579	0.332
Ex 4				0.622	0.593	0.231
Ex 5					0.818	0.235
Ex 6						0.578
ICC [‡] (95% CI)	0.473 (0.387 to 0.559)					
Bitewing [*]						
Ex 1	0.717	0.345	0.597	0.722	0.700	0.413
Ex 2		0.781	0.360	0.425	0.325	0.196
Ex 3			0.770	0.620	0.487	0.363
Ex 4				0.800	0.540	0.388
Ex 5					0.581	0.464
Ex 6						0.622
ICC [‡] (95% CI)	0.538 (0.465 to 0.614)					

[†] Digital subtraction radiography

[‡] Intraclass correlation coefficient

^{*} Weighted Kappa values. Shaded cells correspond to intra-rater reproducibility and unshaded cells are related to inter-rater reproducibility

ICC = Intraclass correlation coefficient; 95% CI = 95% confidence interval

Mean (standard deviation) of Kappa values:

Inter-rater reproducibility

DSR = 0.437 (0.162); Bitewing = 0.463 (0.148); p = 0.544

Intra-rater reproducibility

DSR = 0.654 (0.092); Bitewing = 0.712 (0.081); p = 0.403