DOI: 10.4274/TurkJOrthod.2023.2022.58

TURKISH JOURNAL OF
ORTHODONTICS

Original Article

Inter-examiner Reliability of Two Methods for Scoring Post-Orthodontic White Spot Lesions from Digital Photographs

Aikaterini Papadimitriou[1]*, Diana Mortensen[2]*, Iosif Sifakakis[3], Olympia Salamara[4], Svante Twetman[2], Sotiria Gizani[4]

[1]Aristotle University of Thessaloniki School of Health Sciences, Faculty of Dentistry, Department of Orthodontics, Thessaloniki, Greece
[2]University of Copenhagen Faculty of Health and Medical Sciences, Department of Odontology, Copenhagen, Denmark
[3]National and Kapodistrian University of Athens School of Dentistry, Department of Orthodontics, Athens, Greece
[4]National and Kapodistrian University of Athens School of Dentistry, Department of Paediatric Dentistry, Athens, Greece

Cite this article as: Papadimitriou A, Mortensen D, Sifakakis I, Salamara O, Twetman S, Gizani S. Inter-examiner Reliability of Two Methods for Scoring Post-orthodontic White Spot Lesions from Digital Photographs. *Turk J Orthod.* 2024; 37(2): 79-83

Main Points

- There was a moderate to strong positive relationship between the two systems, and the reliability ranged from moderate to excellent, depending on the examiner.
- For both scoring systems, the agreement concerning sound surfaces and the most severe lesions was high, whereas uncertainties occurred for the less severe white spot lesions (WSLs).
- It was practical to score WSLs in a standardized way from photographs.

ABSTRACT

Objective: To compare the reliability of two scoring systems for detecting white spot lesions (WSLs) from clinical photographs captured during debonding of fixed orthodontic appliances.

Methods: Digital images of 58 healthy adolescents (34 females and 24 males) were examined, depicting 384 buccal surfaces of maxillary incisors, canines, and first premolars. Three trained examiners (E1, E2 and E3) independently evaluated the fully anonymized photos in a randomized order using the Gorelick index (GI) and the modified International Caries Detection and Assessment System (ICDAS II). A 1-2-week interval separated the scorings. Spearman's rank correlation coefficient, Fisher's z-test, and the interclass correlation coefficient (ICC) were applied to compare the scoring methods and express examiner agreement.

Results: The two scoring systems showed a moderate to strong positive relationship, but inter-examiner variations were significant ($p<0.05$). We found moderate to good reliability (ICC 0.60 to 0.84) with the ICDAS II system and good to excellent values with the GI (ICC 0.72 to 0.94), depending on the examiner. The agreement concerning the sound surfaces and the most severe WSLs was perfect, whereas the scoring of the milder lesion stages appeared more uncertain.

Conclusion: A moderate to strong positive relationship was demonstrated between the two methods when scoring the presence and severity of WSLs from digital images. Significant inter-examiner variations affected reliability.

Keywords: Adolescents, caries index, white spot lesions, fixed appliances, orthodontics

*The authors Mortensen and Papadimitriou had equal distribution at the paper.

Corresponding author: Sotiria Gizani, **e-mail:** stgizani@dent.uoa.gr

Received: May 27, 2022 **Accepted:** June 04, 2023 **Publication Date:** June 30, 2024

INTRODUCTION

White spot lesions (WSLs), which are also known as enamel demineralization, are a common and unwanted side effect of treatment with fixed orthodontic appliances.[1,2] The build-up of a dental biofilm (plaque) adjacent to the devices affects oral hygiene and increases the risk of enamel mineral loss.[3] The prevalence of WSLs after orthodontic treatment varies from 2% to 96%, depending on the method and criteria for lesion detection and patients' compliance with recommended preventive measures.[4] The detection and scoring of WSLs rely on clinical visual methods and/or adjunctive technologies such as laser fluorescence, quantitative light-induced fluorescence, and impedance spectroscopy.[5-8] Because clinical scoring is inexpensive, visual inspection of clean, dry tooth surfaces remains the standard care in early lesion detection.[9,10] However, multicenter studies and unforeseen events, like the coronavirus disease-2019 pandemic, may impede access to clinical inspection by calibrated examiners. Thus, scoring WSLs from clinical digital photographs has emerged as a practical and timesaving option.[11,12] Common scoring systems in orthodontic care are Gorelick index (GI)[5] and the International Caries Detection and Assessment System (ICDAS II).[10] To the best of our knowledge, there is a lack of information on the utility of the abovementioned methods when scored from digital images. Therefore, the aim of the present study was to compare the reliability of the Gorelick and ICDAS II indices in scoring WSLs with the aid of clinical digital photographs captured immediately after removal of the fixed orthodontic appliances.

METHODS

Study Design

This retrospective study reevaluated clinical photos from participants in two earlier studies.[8,12] The pooled study group consisted of 58 healthy adolescents (34 females and 24 males) and included 384 buccal surfaces of maxillary incisors, canines, and first premolars. All patients received treatment with fixed orthodontic appliances at the School of Dentistry, National and Kapodistrian University of Athens. The patients were consecutively enrolled, and the inclusion criterion was at least two buccal WSLs at the time of debonding. The exclusion criteria were severe chronic diseases and regular use of xenogenic drugs. The patients and their parents provided written informed consent for the study, and the protocol was approved by the Ethics Committee of the Dental School, National and Kapodistrian University of Athens (approval no.: 409, date: 24.10.2016).

Clinical Procedures

After debonding, the remaining composite material on the buccal tooth surfaces was thoroughly removed with a slow rotating carbide bur, followed by polishing with a rubber cup and pumice paste. After drying with compressed air, three digital photographs (frontal, right, and left lateral) of each patient were obtained with a digital single-lens reflex camera (Nikon D7100 body with a Nikon AF-S VR Micro-NIKKOR 105 mm f/2.8G IF-ED lens) equipped with a polarized filter and a dual flash. One single investigator took all photographs to standardize the quality. The camera was angled around 20° perpendicular to the buccal tooth surfaces to minimize flash reflection. The photographs were then anonymized and provided with a specific research code. Three trained examiners (E1, E2 and E3) independently evaluated the photos in a randomized order using a high-definition screen in a darkened room. The first session included the assessment with the GI (Score 1= no visible white spot or surface demineralization; Score 2= WSL covered less than one-third of the tooth surface, no surface disruption; Score 3= WSL covered more than one-third of the surface, with roughened surface; and Score 4= visible cavitation).[5] After 1-2 weeks, the same three examiners reassessed the photos in a blinded manner with the merged ICDAS II index;[7,11] Score 0a= no visible signs of demineralization (ICDAS 0); Score 1a= enamel caries when viewed dry or wet (ICDAS 1 and 2); Score 2a= localized enamel breakdown or underlying dark shadow (ICDAS 3 and 4); Score 3a= dentin caries with visible cavity (ICDAS 5). Figure 1 depicts ICDAS II score 1a lesions, while Figure 2 showcases examples of ICDAS II score 2a lesions. The three examiners undertook a consensus-based training program with both methods before the original studies.

Statistical Analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS Inc, version 26.0, Chicago, IL, USA). We assumed that the four categories of the

Figure 1. Representative image of the ICDAS II score 1a lesions
ICDAS II, International Caries Detection and Assessment System

Figure 2. Representative image of the ICDAS II score 2a lesions
ICDAS II, International Caries Detection and Assessment System

scoring systems would correspond to each other. The normality of the distribution was checked before the parametric tests were applied. Spearman´s rank correlation coefficients for the two scoring methods were calculated, and the obtained coefficients were compared using Fisher's *z*-test. The difference between the paired proportions of sound surfaces vs. surfaces with WSL was tested for each examiner using the McNemar test. The distribution of scores within the classification systems was compared using chi-squared tests. The interclass correlation coefficient (ICC) was used to assess examiner agreements. P-values <0.05 were considered statistically significant.

RESULTS

This study included a total of 384 buccal sites from 58 patients. The mean age was 15 years, ranged from 13.1 to 17.1 years. Table 1A-C shows the cross-tabulation of the two scoring methods by the three examiners. The Spearman correlation coefficients for E1, E2, and E3 were 0.53, 0.80, and 0.61, respectively. This coefficient was significantly higher (p<0.05) for E2 than for E1 and E3. When dichotomized to sound surfaces vs. surfaces with WSL, only examiner E3 demonstrated a significant difference between the proportions obtained from the two methods (Table 2). Examiner E2 scored significantly higher WSL levels (p<0.05) with the modified ICDAS II system than the other two examiners, but no such differences were found with the GI. The ICC values are presented in Table 3. The values ranged from 0.60 to 0.84 with the ICDAS II system, indicating moderate to good reliability. The corresponding figures for the GI varied between 0.72 and 0.94, which suggested good to excellent reliability.

Table 1A-C. Cross-tabulation of the WSL scores registered by the three examiners (Table 1A for examiner E1, 1B for E2 and 1C for E3) with the modified ICDAS II index and the GI

Table 1A.

	GI					
ICDAS II	**1**	**2**	**3**	**4**	**Total**	
0a	57	25	5	0	87	22.7%
1a	30	187	33	0	250	65.1%
2a	3	14	26	1	44	11.5%
3a	1	0	0	2	3	0.8%
Total	91	226	64	3	384	100.0%
	23.7%	58.9%	16.7%	0.8%	100.0%	

Table 1B.

	GI					
ICDAS II	**1**	**2**	**3**	**4**	**Total**	
0a	50	0	0	0	50	13.0%
1a	0	274	46	0	320	83.3%
2a	0	0	8	3	11	2.9%
3a	0	0	0	3	3	0.8%
Total	50	274	54	6	384	100.0%
	13.0%	71.4%	14.1%	1.6%	100.0%	

Table 1C.

	GI					
ICDASII	**1**	**2**	**3**	**4**	**Total**	
0a	61	17	0	0	78	20.3%
1a	33	202	34	0	269	70.1%
2a	1	17	17	1	36	9.4%
3a	0	0	0	1	1	0.3%
Total	95	236	51	2	384	100%
	24.7%	61.5%	13.3%	0.5%	100.0%	

WSL, white spot lesion, ICDAS II, International Caries Detection and Assessment System, GI, Gorelick index

DISCUSSION

This is the first study to compare the reliability of two scoring systems for the presence of WSLs immediately after the debonding of fixed orthodontic appliances. Instead of visual clinical inspection, three digital high-resolution images were evaluated in each patient to study the buccal surfaces of the maxillary incisors, canines, and first premolars. The examiners found scoring from photographs practical and time-saving because the assessments could be performed outside regular office hours and in a standardized mode. The scoring index developed by Gorelick et al.[5] was exclusively developed for orthodontic patients, while the ICDAS II index was an assessment system for coronal caries,[7] which was later adapted for orthodontic patients.[11] However, it should be noted that the two systems might not be directly comparable; the practical difference between these methods is that the former focuses on lesion extension (surface area) and the latter on lesion discrimination. However, the most advanced stages with both systems clearly denote cavitation.

The main findings of this study indicated moderate to excellent concordance between the two scoring systems, depending on the examiner. Although the agreement concerning the sound

Table 2. Difference between the paired proportions of sound surfaces vs. presence of WSL for the two scoring systems by examiner

Examiner	Difference, %	95% confidence interval	p-value
E1	1.04	-3.04 to 5.12	0.71*
E2	0.00	-0.72 to 0.72	1.00**
E3	4.43	0.81 to 8.05	0.02***

Statistical significance: *p<0.05, **p<0.01, ***p<0.001
WSL, white spot lesion

Table 3. ICC for the two WSL scoring methods and by examiner (E1, E2 and E3). ICC values between 0.5 and 0.75 indicate a moderate reliability, between 0.75 and 0.90 a good reliability, and values greater than 0.90 indicate an excellent reliability

	E1 vs. E2	E1 vs. E3	E2 vs. E3
ICDAS II	0.60	0.84	0.74
GI	0.74	0.94	0.72

ICC, interclass correlation coefficient, WSL, white spot lesion, ICDAS II, International Caries Detection and Assessment System, GI, Gorelick index

surfaces and the most severe lesions was perfect, uncertainties occurred for the lower scores with both indices. For example, the prevalence of Gorelick score 2 varied between 59% and 71% among the examiners, and the values for ICDAS II score 1a ranged from 65% to 83%. Over time, however, this might not be a major problem because the prevalence of minor post-WSLs (GI Score 2) seems to drop by over 50% one year after debonding due to natural remineralization and secondary prevention.[1,13] However, inter-examiner variability may influence the estimated prevalence of post-orthodontic WSLs in clinical trials. In the present study population, the prevalence would have ranged from 75% to 87%, with no major differences between the scoring methods. This prevalence of WSLs was indeed higher than expected in an "average" population of orthodontic patients,[14,15] but was explained by the inclusion criteria, in which only patients with WSLs were enrolled.

The present study indicated that both scoring methods may be useful in the clinic, but it was not possible to argue in favor of one method over the other, due to a lack of formal validation.[16,17] Obviously, this was not within the scope of this project, as a validation study necessitates clinical access to patients and adherence to a predetermined standard. In clinical research, the use of digital images offers several advantages, such as the possibility of masking patients and enabling a random order of examination and reassessment of previously collected study groups. It also facilitates the performance of multicenter studies, which are often necessary to recruit a sample size with sufficient power. The impact of inter-examiner variability can also be limited by involving multiple independent examiners.

Study Limitations

To minimize the photographic shortcomings, a standardized exposure procedure was used, and the camera was equipped with a polarizing filter and angled to avoid flash reflections from the tooth surfaces, to mimic enamel demineralization. The three examiners went through a consensus-based training program before the evaluation sessions; however, it is possible that further education and experience could have improved the concordance. No dropouts of patients or images due to technical errors were present because this study was a reexamination of material from two previous trials. It is however important to note that the sample size was relatively small, and further research with a larger study population would provide more robust evidence for the reliability of the two methods. Another limitation of the study was the use of a slow rotating carbide bur to remove the remaining composite material after bracket removal, a procedure that undoubtedly affected the enamel surface and potentially influenced the subsequent scorings.

CONCLUSION

A moderate to strong positive relationship was found between the two methods for scoring the presence and severity of WSL development from clinical photographs, which were exposed immediately after debonding of fixed orthodontic appliances. Significant inter-examiner variations were obtained; however, the agreement-concerning the sound surfaces and the most severe WSLs was high. Clinicians involved in practice-based research might therefore undergo structured training to visually classify WSLs using any of the two scoring systems to improve the reliability and quality of the outcome measure.

Ethics

Ethics Committee Approval: The research protocol was submitted on 11/07/2016 and approved on 24/10/2016 by the Ethics Committee of the Dental School, National and Kapodistrian University of Athens, according to Helsinki's Declaration (approval no.: 409, date: 24.10.2016).

Informed Consent: The patients and their parents provided written informed consent for the study.

Author Contributions: Concept - I.S., S.T., S.G.; Design - I.S., S.T., S.G.; Supervision - I.S., S.T., S.G.; Materials - A.P., D.M., I.S., O.S.; Data Collection and/or Processing - A.P., D.M., O.S.; Analysis and/or Interpretation - D.M.; Literature Review - I.S., S.T. S.G.; Writing - A.P., D.M., S.T.; Critical Review - I.S., S.T. S.G.

Declaration of Interests: All authors declare that they have no conflict of interest.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Willmot D. White spot lesions after orthodontic treatment. *Semin Orthod*. 2008;14(3):209-219. [CrossRef]
2. Benson PE, Parkin N, Dyer F, Millett DT, Germain P. Fluorides for preventing early tooth decay (demineralised lesions) during fixed brace treatment. *Cochrane Database Syst Rev*. 2019;2019(11):CD003809. [CrossRef]
3. Tufekci E, Dixon JS, Gunsolley JC, Lindauer SJ. Prevalence of white spot lesions during orthodontic treatment with fixed appliances. *Angle Orthod*. 2011;81(2):206-210. [CrossRef]
4. Sonesson M, Bergstrand F, Gizani S, Twetman S. Management of post-orthodontic white spot lesions: an updated systematic review. *Eur J Orthod*. 2017;39(2):116-121. [CrossRef]
5. Gorelick L, Geiger AM, Gwinnett AJ. Incidence of white spot formation after bonding and banding. *Am J Orthod*. 1982;81(2):93-98. [CrossRef]
6. Beerens MW, Boekitwetan F, van der Veen MH, ten Cate JM. White spot lesions after orthodontic treatment assessed by clinical photographs and by quantitative light-induced fluorescence imaging; a retrospective study. *Acta Odontol Scand*. 2015;73(6):441-446. [CrossRef]
7. Almosa NA, Lundgren T, Aldrees AM, Birkhed D, Kjellberg H. Diagnosing the severity of buccal caries lesions in governmental and private orthodontic patients at debonding, using the ICDAS-II and the DIAGNOdent Pen. *Angle Orthod*. 2014;84(3):430-436. [CrossRef]
8. Mortensen D, Gizani S, Salamara O, Sifakakis I, Twetman S. Monitoring regression of post-orthodontic lesions with impedance spectroscopy: a pilot study. *Eur J Orthod*. 2019;41(4):415-419. [CrossRef]
9. Pretty A, Ekstrand KR. Detection and monitoring of early caries lesions: a review. *Eur Archs Paediatr Dent*. 2016;17:13-25. [CrossRef]

10. Ekstrand KR, Gimenez T, Ferreira FR, Mendes FM, Braga MM. The International Caries Detection and Assessment System - ICDAS: A Systematic Review. *Caries Res.* 2018;52(5):406-419. [CrossRef]
11. Almosa NA, Lundgren T, Bresin A, Birkhed D, Kjellberg H. Diagnosing the severity of buccal caries lesions in orthodontic patients at de-bonding using digital photographs. *Acta Odontol Scand.* 2014;72(6):474-480. [CrossRef]
12. Salamara O, Papadimitriou A, Mortensen D, Twetman S, Koletsi D, Gizani S. Effect of fluoride varnish with functionalized tri-calcium phosphate on post-orthodontic white spot lesions: an investigator-blinded controlled trial. *Quintessence Int.* 2020;51(10):854-862. [CrossRef]
13. Sonesson M, Brechter A, Lindman R, Abdulraheem S, Twetman S. Fluoride varnish for white spot lesion prevention during orthodontic treatment: results of a randomized controlled trial 1 year after debonding. *Eur J Orthod.* 2021;43(4):473-477. [CrossRef]
14. Sonesson M, Brechter A, Abdulraheem S, Lindman R, Twetman S. Fluoride varnish for the prevention of white spot lesions during orthodontic treatment with fixed appliances: a randomized controlled trial. *Eur J Orthod.* 2020;42(3):326-330. [CrossRef]
15. Sonesson M, Twetman S, Bondemark L. Effectiveness of high-fluoride toothpaste on enamel demineralization during orthodontic treatment-a multicenter randomized controlled trial. *Eur J Orthod.* 2014;36(6):678-682. [CrossRef]
16. Benson P. Evaluation of White Spot Lesions on Teeth with Orthodontic Brackets. *Seminars in Orthodontics*. 2008;14(3):200-208. [CrossRef]
17. Huang J, Chiang B, Mills B, et al. Effectiveness of MI Paste Plus and PreviDent fluoride varnish for treatment of white spot lesions: a randomized controlled trial. *Am J Orthod Dentofacial Orthop.* 2013;143(1):31-41. [CrossRef]