



Reliability assessment of the classification for facial peri-implant soft tissue dehiscence/deficiencies (PSTDs): A multi-center inter-rater agreement study of different skill-level practitioners

Shayan Barootchi^{1,2} | Leonardo Mancini³ | Teresa Heck¹ |
Giovanni Zucchelli^{1,4} | Martina Stefanini⁴ | Emilia Kazarian⁵ |
Giulio Rasperini^{1,6} | Hom-Lay Wang¹ | Lorenzo Tavelli^{1,2,7}

¹ Department of Periodontics & Oral Medicine, University of Michigan School of Dentistry, Ann Arbor, Michigan, USA

² Center for clinical Research and evidence synthesis In oral TissuE RegeneratIOn (CRITERION), Ann Arbor, Michigan, USA

³ Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

⁴ Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy

⁵ Private practice, Moscow, Russia and Piacenza, Italy

⁶ Department of Biomedical, Surgical and Dental Sciences, Foundation IRCCS Ca' Granda Polyclinic, University of Milan, Milan, Italy

⁷ Department of Oral Medicine, Infection, and Immunity, Division of Periodontology, Harvard School of Dental Medicine, Boston, Massachusetts, USA

Correspondence

Shayan Barootchi, Department of Periodontics and Oral Medicine, University of Michigan, School of Dentistry, 1011 N. University Avenue, Ann Arbor, MI 48109 USA.

Email: shbaroot@umich.edu

Abstract

Background: Along with the popularity of dental implants, implant esthetic complications are also on the rise. Recently a classification was proposed to comprehensively evaluate these conditions, with the definition of peri-implant soft tissue dehiscence/deficiencies (PSTDs). The aim of this article was therefore to test the inter-examiner agreement when utilizing the established rubrics among 25 standardized cases and 34 clinicians of different skill levels.

Methods: Twelve periodontal residents, 12 general dentists, and 10 periodontists participated in this study. All examiners were provided with photographs of 25 single PSTDs and asked to rate all cases based on the proposed classification at a single timepoint. Variance components analysis was conducted with multilevel regression fit in a Bayesian framework to obtain uncertainty intervals for fractional variance contributions and interclass correlation values (ICC) to assess the agreement in the rating of all cases, among all examiners, different skill-level practitioners, and to compare their responses relative to the judgment of a gold standard examiner.

Results: Overall, the results showed reproducible and consistent responses among the 34 examiners, and in each subgroup of skill-level, comparable to that of the gold examiner. Nevertheless, periodontists and residents were more likely to agree with the response of the gold standard examiner in their assessments of class and subclass of the PSTDs.

Conclusions: The proposed PSTD classification showed reproducible assessments among all examiners, and between examiners of the same skill-level. The response of the gold standard examiner was more in line with the assessment of the periodontists and periodontal residents.

KEYWORDS

dental implants, diagnosis, esthetics, evidence-based dentistry, gingival recession, periodontics



1 | INTRODUCTION

Dental implants have become a popular treatment modality for the replacement of missing teeth, particularly in the recent years.¹⁻³ The tremendous rise in the popularity of implants has also accompanied certain complications and adverse events.⁴⁻⁷ Among them, esthetic challenges have rapidly become an emerging concern for clinicians and patients, especially given the rise in patients' esthetic demands.^{8,9}

Implant esthetic complications can include discoloration of the peri-implant soft tissues, apical shifting of the implant mucosal margin and exposure of the implant components, an implant-supported crown that is longer than its homologous contralateral tooth, or a combination of these events. Throughout the literature, and along with the increased prevalence of these concerns, different terminologies have been utilized for their description, such as "mucosal recession," "soft tissue dehiscence," "midfacial recession," "marginal soft tissue recession," "soft tissue defect," and so on.¹⁰⁻¹⁶ In addition, for several years their incidence has been reported without a uniform or standard definition, and although studies have described surgical procedures for the correction of these adverse events, the lack of a uniform diagnostic system has prevented comparisons among the proposed treatments options, or the outcomes of therapy.^{11,14,16-18} A soft tissue dehiscence has occasionally been described as the exposure of the prosthetic abutment or implant neck,^{12,15,19} whereas others have used the homologous contralateral tooth as reference, thus considering a visibly longer implant crown as an esthetic concern.^{11,16,20} Furthermore, clinical studies investigating the treatment of implant esthetic complications often lack information of the position of the dental implant relative to the bony housing (e.g., if buccally displaced) and the height of the interproximal soft tissues, which are among the critical parameters affecting the treatment outcomes.¹⁵

Our group recently proposed a classification system for the description of these esthetic challenges, defining them collectively at clinically healthy implant sites, as peri-implant soft tissue dehiscence/deficiencies (PSTDs).²¹ The goal of this classification was to formulate a uniform and objective diagnostic tool for clinicians and researchers and pave the way for providing recommendations towards treatment options, facilitating the decision-making process based on the specific type (class and subclass) of the PSTDs. Nevertheless, as with any classification or grading system, it is vital to assess the consistency of the application of its proposed rubric. Therefore, the aim of the present study was to assess the reliability in the application of this classification, for classifying PSTDs at single implant sites

in the esthetic zone across a diverse range of dental practitioners with different skill levels and expertise.

2 | MATERIALS AND METHODS

2.1 | Study design and registration

The present study was designed to test the reliability of the classification for facial PSTDs at single implant sites,²¹ and the variation in the responses of different skill-level classes of practitioners, and lastly their comparison to that of a gold standard examiner.

Twenty-five standardized photographs of a variety of PSTD classes and subclasses were utilized for this study. The patients whose cases were utilized had provided their informed consent, and the study protocol was approved by the University of Michigan Medical School Institutional Review Board (IRBMED, HUM00176741), in accordance with the Declaration of Helsinki of 1975, revised in Tokyo in 2013. The manuscript is also prepared following the EQUATOR guidelines Standards for Reporting Qualitative Research (SRQR).²²

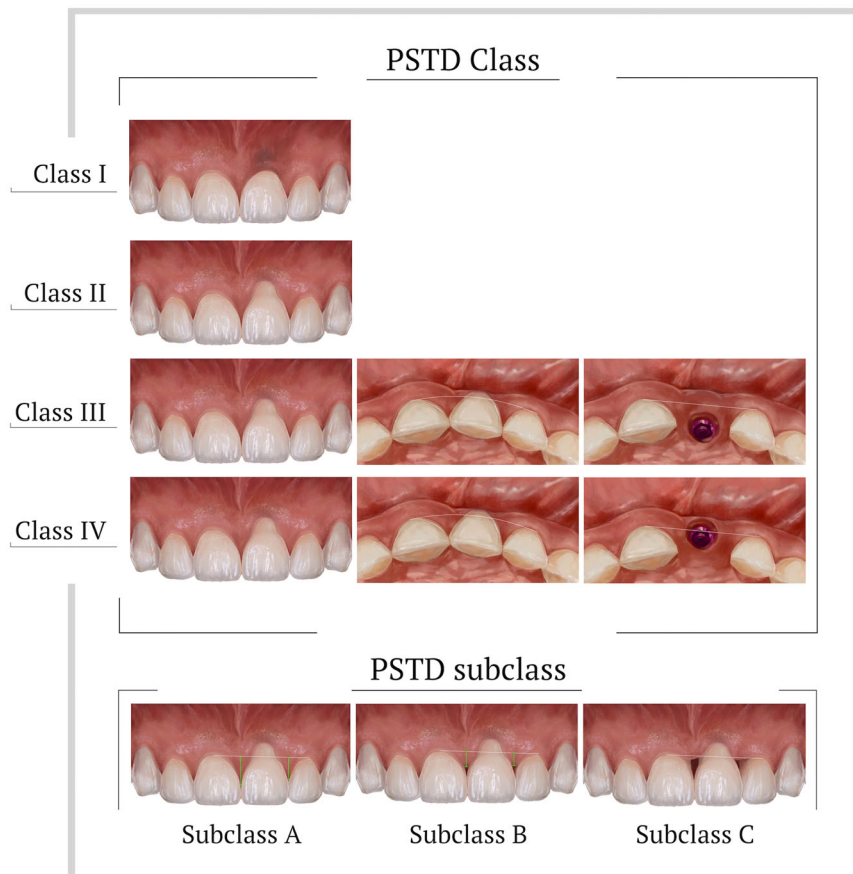
2.2 | Classification of facial PSTD

The classification on the types (classes and subclasses) of PSTDs is explained in detail in the original report.²¹ In summary, a PSTD was defined as any complication which can manifest as either an apparent mid-facial mucosal recession (apical shift of the peri-implant mucosal margin), the display of the implant component(s) through the mucosa (such as a noticeable appearance of the grayish color of the implant abutment), or even a discrepancy in the length of an implant crown relative to its homologous natural tooth.²¹ Based on the position of the soft tissue (mucosal) margin of the implant-supported crown in relation to the gingival margin of the homologous natural tooth, a PSTD can be categorized into four classes (I, II, III, or IV), and further grouped into three subclasses (a, b, or c) depending on the bucco-lingual position of the implant fixture, and on the height of the anatomical papillae, respectively.²¹ Figure 1 summarizes the above-mentioned classification.

2.3 | Assessment of inter-rater reliability of the classification for PSTD

A total of 34 practitioners from different centers and different backgrounds/skill-levels participated in this relia-

FIGURE 1 Schematic representation to summarize the classification on peri-implant soft tissue dehiscence/deficiencies (PSTD)²¹



bility study, including 12 general dentists (from a single center, Department of Life, Health and Environmental Sciences, University of L'Aquila), 12 current periodontal residents (from the University of Michigan Department of Periodontics and Oral Medicine, Ann Arbor, Michigan, USA), and 10 American Academy of Periodontology board-certified periodontists (based in Michigan, USA). All examiners were similarly provided with the full text of the above-mentioned main publication explaining the PSTD classification.²¹

Next, the examiners were provided with a series of slides (each slide containing a frontal and an occlusal shot), each corresponding to a single case of a PSTD (25 in total) along with its designated case number, compiled without a specific order into a single document (Adobe Acrobat Pro DC for Mac©, version 2021.005.20058). The examiners were also distributed a pre-designed spreadsheet containing the same corresponding case numbers per row, and the two columns of "Class," and "Subclass," which required their response according to the afore-mentioned criteria of the classification. Prior to the assessment of the distributed cases, the examiners were provided with clarification with regards to their task, or the rubrics of the classification system, if they inquired.

All clinical photographs had been obtained by a single photographer (L.T.), under standard conditions and a shooting protocol (same setting and with a single camera; full-frame digital single lens-reflex (DSLR) camera,[#] with electro-focus (EF) 100 mm f/2.8L Macro Lens,^{**} and Macro Ring flash).^{††} Meticulous care was taken during the compilation and distribution of cases to maintain the original standardized 1:1 ratio of all clinical photos, without adjustments to the quality of the pictures (brightness, contrast, color and so on). None of the examiners had any prior knowledge of the provided cases, and in order to further reduce bias, no other information or guidance was provided to the examiners, upon the initiation of their assessment.

Ample time was provided to all examiners to accomplish this task without any time limitation. The preparations and distribution of all cases were performed by a single study team member who did not take part in the reliability assessment (L.T.).

[#] Canon EOS 5D Mark IV, Canon, Ota City, Tokyo, Japan

^{**} Canon, Ota City, Tokyo, Japan

^{††} MR-14EX II Macro Ring Lite, Canon, Ota City, Tokyo, Japan



2.4 | Inclusion of a gold standard examiner

A separate examiner, with knowledge of all 25 cases and expertise in periodontal and peri-implant plastic surgery (G.Z.) who was part of the original conception of the PSTD classification, but did not take part in the reliability assessment, also rated all cases, to provide for the “true” response of all cases. The aim was to assess whether a set of skill levels were more likely to agree or disagree with the judgment of the gold standard, than others.

2.5 | Statistical analysis and inter-rater agreement assessment

The gathered responses were first descriptively assessed for crude agreement between raters. Next, similar to previous methodology,⁹ logistic variance components analysis was used to decompose the variation in the individual parameters of the PSTD score into variance contributions for cases and for examiners. The analysis was conducted using multilevel generalized linear regression, fit in a Bayesian framework to facilitate construction of uncertainty intervals for fractional variance contributions and intraclass correlation values. The analysis produces estimates of the variance explained by cases, examiners, and the variance that is unexplained. The case variance reflects true differences among the cases. The examiner variance reflects systematic differences among the examiners in item-level responses, for example, if it is inflated by an examiner who is consistently more prone to endorsing certain items than the others.

The intraclass correlation coefficient (ICC) was defined by dividing the variance from the cases by the sum of variances from all sources. It is desirable for this ICC to be high.^{23–25} The variance contribution for examiners should also be small, because it reflects systematic differences in rating behaviors by different examiners. The analysis was conducted using a Bayesian approach to obtain 95% credible intervals (CI), with *Z*-scores to convey statistical significance.

To assess agreement with the gold standard examiner, a logistic variance components model was fit with “agreement” (Yes/No) as the dependent variable, random effects for case and examiners, and fixed-effects for skill levels.

The analysis was performed by an author with experience in statistical analysis (S.B.) who had not taken part in the reliability assessment, using a designated software^{‡‡} and the following packages lme4,²⁶ lmerTest,²⁷ Rcpp,^{28–30}

TABLE 1 Inter-rater agreement as measured by the intraclass correlation coefficient (ICC) for each PSTD class and subclass, for all raters of all skill levels

Category	ICC	95% CIs (lower, upper bound)
PSTD class		
I	0.907	0.718, 0.993
II	0.992	0.963, 0.999
III	0.995	0.975, 0.999
IV	0.995	0.978, 0.999
III/IV	0.993	0.973, 0.999
PSTD subclass		
A	0.947	0.805, 0.999
B	0.985	0.952, 0.998
C	0.996	0.979, 0.999

Abbreviations: ICC, intraclass correlation coefficient; CI, credible intervals; PSTD, peri-implant soft tissue dehiscence/deficiencies.

brms,^{31,32} arm,³³ tidy,³³ and tidybayes.³⁴ Box plots were produced using the ggplot2 package.³⁵

3 | RESULTS

3.1 | General characteristics of the examined PSTDs and the overall reliability of the classification

Twenty-five cases of PSTDs (each in a single patient) were assessed by all 34 examiners at a single point. Sixteen PSTDs were in the maxilla, whereas the remaining nine cases were in the mandible. Ten implants with PSTDs were in the lateral incisor position, seven in the central incisor position, five in the premolar area and three at the canine area. Twelve PSTDs were characterized by having an implant-supported crown longer than the homologous tooth and an exposure of the abutment or implant fixture at the same time. Nine implants displayed an implant-supported crown longer than the homologous tooth (without exposure of the abutment/implant fixture), whereas four PSTDs were characterized by an exposure of the abutment or implant fixture with an adequate length of the implant-supported crown.

Table 1 shows the inter-rater agreement for the examined cases per each class and subclass of the PSTD classification. Overall, the classification showed consistency across all practitioners in its use relative to both the assessment of class (ICC scores of ≈ 0.9), and subclass (ICC scores of ≈ 0.95).

‡‡ RStudio Version 1.3.959, RStudio, Inc., Boston, Massachusetts

TABLE 2 Inter-rater agreement as measured by the intraclass correlation coefficient (ICC) for each PSTD class and subclass, for raters of each skill level

Category	General practitioners	Periodontal residents	Periodontists
	ICC (95% CIs)	ICC (95% CIs)	ICC (95% CIs)
PSTD class			
I	0.98 (0.92, 0.99)	0.99 (0.95, 0.99)	0.96 (0.81, 0.99)
II	0.98 (0.90, 0.99)	0.97 (0.88, 0.99)	0.98 (0.93, 0.99)
III	0.97 (0.87, 0.99)	0.97 (0.88, 0.99)	0.97 (0.87, 0.99)
IV	0.99 (0.95, 0.99)	0.98 (0.93, 0.99)	0.99 (0.94, 0.99)
III/IV	0.99 (0.97, 0.99)	0.98 (0.91, 0.99)	0.98 (0.94, 0.99)
PSTD subclass			
A	0.92 (0.69, 0.99)	0.95 (0.77, 0.99)	0.96 (0.81, 0.99)
B	0.98 (0.94, 0.99)	0.98 (0.93, 0.99)	0.99 (0.97, 0.99)
C	0.99 (0.97, 0.99)	0.99 (0.95, 0.99)	0.99 (0.97, 0.99)

Abbreviations: CI, credible intervals; ICC, intraclass correlation coefficient; PSTD, peri-implant soft tissue dehiscence/deficiencies.

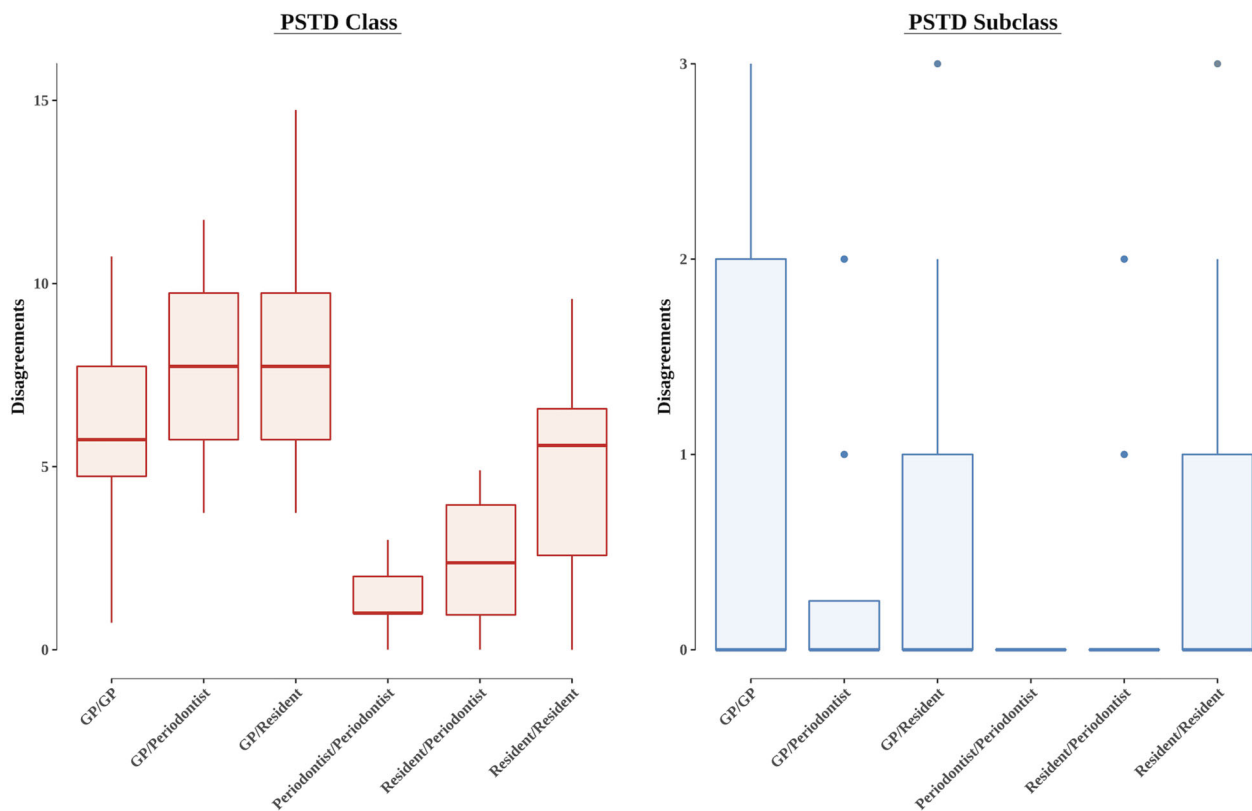


FIGURE 2 Distribution of rater-pair disagreements by skill level for the 25 cases of peri-implant soft tissue dehiscence/deficiencies (PSTD) classes and subclasses. GP, general practitioners

3.2 | Skill-level subgroup assessment of the reliability in usage of the classification

Table 2 shows the inter-rater agreement, in terms of ICC for each of the three skill-level categories, and Figure 2 displays the distribution of the number of disagreements for

each pair of skill levels for PSTD class and subclass. All examiners in their own categories seemed to agree on the assessment of the cases. All ICC scores were above 0.9 in all skill-level categories, demonstrating a high level of agreement among all practitioners when using the PSTD classification.

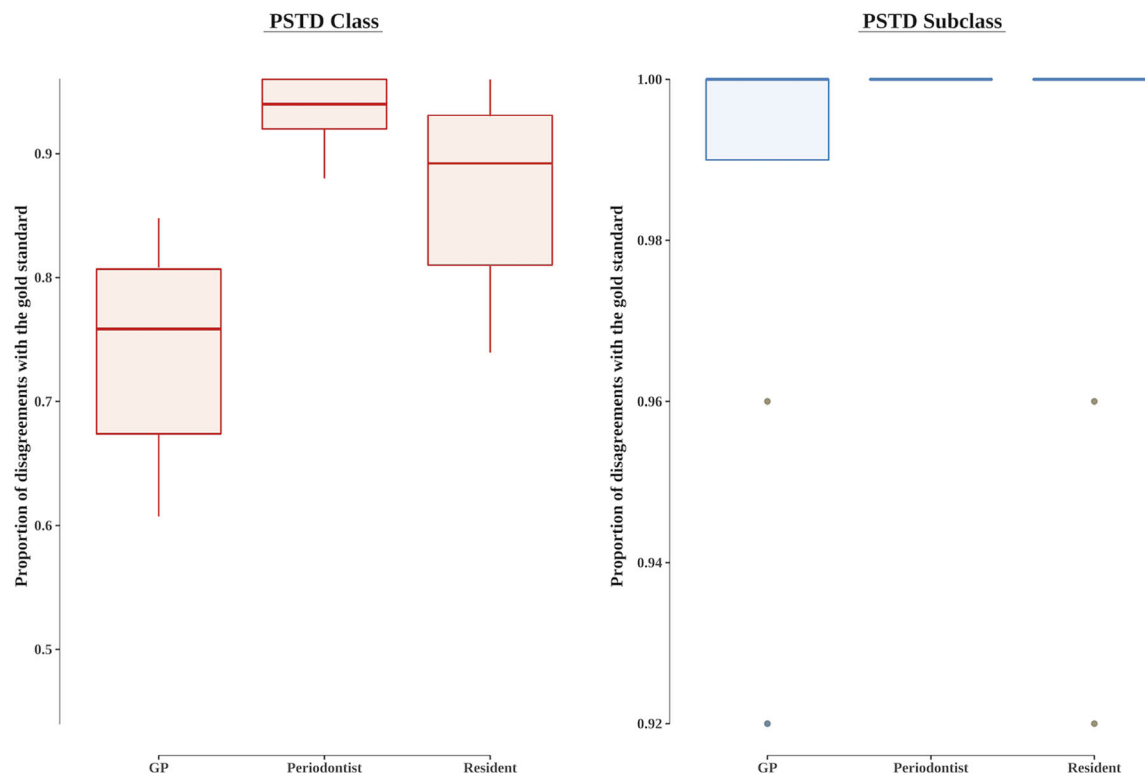


FIGURE 3 Box plots presenting the proportion of the agreement for the examined cases relative to the judgment of the gold standard examiner. PSTD, peri-implant soft tissue dehiscence/deficiencies; GP, general practitioners

3.3 | Assessment of consistency in response of examiners with that of the gold standard

For each skill level, we began by calculating the crude proportion of disagreements with the gold standard. Figure 3 shows the distribution of the percentage of the rate of agreement with the gold standard examiner, for PSTD class and subclass, per skill-level category of clinicians. For the assessment of the PSTD classes, from the total 850 observations (34 examiners for 25 cases), 76.6% of the responses of the general practitioners, 85% of the responses of periodontal residents, and 93.2% of the responses of the periodontists agreed with the rating of the gold standard examiner (Figure 3).

Table 3's top panel presents the results of the analysis of agreement in PSTD classes with the gold standard's response, as per skill-level category of examiners, using a binomial mixed model to account for examiner and case effects. Based on this model, we observed that periodontists and periodontal residents were significantly less likely to disagree with the response of the gold standard examiners, than general practitioners. However, there was no statistical significance when comparing the disagreement rates between residents and periodontists. Fur-

ther, based on the variance decomposition for examiners and cases, we observed that 93% of the variance is attributable to cases with the remainder attributable to examiners.

For the PSTD subclasses, the responses were vastly in line with the response of the gold standard examiner. Such that from the 850 observations, there were only eight instances in which there was a disagreement with the response of the gold standard, five of which were by general practitioners (three examiners and five cases), and three which occurred by residents (one error per each of the three residents) (Figure 3).

Thus, as displayed by the results of the model (Table 3), the probability of a periodontist disagreeing with the gold standard (or thereby making an "error") was very small for PSTD subclasses. Additionally, because of the very small number of errors for PSTD subclasses, it was not possible to statistically compare the error rates between skill-levels.

In general, for both PSTD Class and Subclass assessments, it appeared that the highest rate of agreement with the gold standard examiner was obtained by the periodontists, closely followed by the periodontal residents (without significant differences), and then the general practitioners, with statistical differences only for PSTD class, but not subclasses.

**TABLE 3** Model-based assessment of agreement to gold standard, for PSTD class and subclass

Category		Model summaries			
		Estimate	Std. Error	95% CI (LB, UB)	Z-score
Class	Population-level fixed effects				
	Intercept	-0.85	0.58	-2.02, 0.25	-1.46
	Skill-level: periodontist	-3.40	0.50	-4.44, -2.46	-6.8 ^a
	Skill-level: resident	-2.12	0.41	-2.99, -1.4	-5.17 ^a
	Group-level random-effect standard deviations				
	Case	2.46	0.51	1.65, 3.63	4.82
Examiner	0.66	0.21	0.27, 1.09	3.14	
Subclass	Population-level fixed effects				
	Intercept	-9.47	3.58	-18.70, -4.81	-2.64 ^a
	Skill-level: periodontist	-40.37	44.40	-154.92, -2.44	-0.909 ^a
	Skill-level: resident	-0.99	2.67	-6.71, 4.09	-0.37
	Group-level random-effect standard deviations				
	Case	3.03	1.81	0.73, 7.97	1.67
Examiner	3.59	2.41	0.66, 9.59	1.48	

Abbreviations: CI, credible intervals; LB, lower bound; UB, upper bound.

^aDenotes statistical significance.

4 | DISCUSSION

The present study aimed at evaluating the inter-rater reliability in the use of a novel classification for PSTDs among a series of operators with different skill levels and experiences, as well as a comparison of their response to that of the “true” assessment of a gold standard examiner.

Several classification systems have been proposed and validated in the field of periodontology and implant dentistry.^{9,36–38} Relative to gingival recessions in natural dentition, the 2017 World Workshop³⁹ adopted the classification proposed by Cairo and coworkers³⁶ based on the interproximal attachment levels, which yielded a high inter-rater agreement of 0.86 among operators when assessing 25 patients with gingival recessions.³⁶ Similarly, in the present study, we observed ICC scores of ≥ 0.9 for both assessments of PSTD class and subclass among all 34 examiners, and individually among each of the three skill-level categories, demonstrating a high level of agreement.²³

The Root Coverage Esthetic Score (RES),^{37,40} which has been routinely utilized throughout the literature,^{41–46} is another validated classification system, used for professional esthetic evaluation of the treatment of gingival recession defects. A study by Isaia et al,⁴⁷ evaluated the inter-rater agreement in the application of RES among residents, faculties, and pre-doctoral dental students in a similar design to the present report. The authors found comparable and high ICC scores among the overall assessment of their three groups of examiners, similar to our research among general practitioners, periodontal residents, and periodontists. A novelty from our study; however, was the

inclusion of a gold standard examiner who had knowledge of all cases, extensive experience in treating PSTDs, and was originally involved in the conception of the utilized PSTD classification.¹⁶ To the best of our knowledge, such comparison to that of a single “true” response, is not commonly performed in reliability assessments in the field of periodontology. The merit of this additional component includes the notion that aside from evaluation of consistency and uniformity among the examiners, and in the application of the proposed rubric, we can also analyze the “correctness” of their response. Which is also one of the reasons why we utilized Bayesian multi-level regression models for the inter-reliability assessment of this study (others being the heterogeneous group of examiners, ability to compare ratings among different classes of examiners, and construction of uncertainty intervals for the ICCs).

Indeed, despite observing an overall high rate of inter-rater agreement in the response of all 34 examiners, and among each skill-level category, we noted that periodontists and periodontal residents were generally more likely to agree with the judgment of the gold standard examiner, which would presumably be the true response. It is reasonable to assume that because the treatment of PSTDs requires a certain level of training in periodontology or knowledge in implant dentistry, periodontists and residents would yield a closer and more correct assessment of the cases. It is also possible that general practitioners are less likely to encounter such conditions because of their wider scope of practice, potentially limiting their attention to the presence such details. On the other hand, periodontal residents and periodontists would be more inclined



to notice these concerns, as it relates to their main area of expertise and clinical practice. For instance, an important factor which can suggest treatment versus removal of an implant with a PSTD, is the bucco-lingual positioning of the implant fixture, which is also the main criteria for determining the Class of the PSTD.⁴⁸ The height of the interproximal soft tissues is another crucial component dictating the surgical approach and the necessity of crown removal.⁴⁹ Therefore, it may be reasonable to assume that both periodontists and periodontal residents would focus more on such details, which essentially defines the class and subclasses of the PSTD classification,¹⁶ and are also the main factors to consider before determining the most suitable approach for the correction of the PSTD.

The proposal of a classification for a pathological or unesthetic condition is certainly not an easy task to accomplish. A classification system should ideally include all possible manifestations of a condition while clustering them under simple and unequivocal categories. Next, we believe that a classification system should also be evaluated in terms of reliability (consistency in use) and clinical significance (therapeutic outcomes). Our group introduced this classification scheme for standardizing the diagnosis of PSTDs, and for providing recommendations for their treatment. The present study showed a high level of agreement among all practitioners when identifying the classes and subclasses of PSTDs, demonstrating consistency in its use when identifying these cases. When comparing the evaluations of all examiners with the gold standard, those with more experience in the field of periodontology and implant dentistry were found to have more similarities in their responses, and to that of the gold standard examiner. This indicates that, although the present classification system was found reliable and reproducible among different examiners, a limited amount of training can be beneficial, for a more accurate assessment of PSTDs, in particular as it relates to their treatment.

Nevertheless, we noticed that the response of periodontal residents to that of board-certified periodontists, and relative to the gold standard examiner did not differ significantly. Thus, even a minimal training or background could be sufficient for an accurate assessment and diagnosis of these conditions. However, in light of the prognostic value of the proposed classification system, future studies are still needed and are currently underway to assess its predictability.

5 | CONCLUSIONS

Within the limitations of the current research, we conclude that the proposed classification for facial PSTDs can

bear reproducible results among a diverse number of clinicians and serve as a reliable tool for evaluating these conditions. However, periodontists and periodontal residents were more likely to agree in their responses, and to the response of the gold standard examiner, for the assessment of the class of PSTD. The application of the introduced classification system can aid in standardized assessment of PSTD, and in a uniform comparison of proposed treatments for these conditions, and therapeutic outcomes in clinical research.

ACKNOWLEDGMENTS

The authors are grateful to all those who participated in this study as examiners and for their valued time and effort in reviewing the cases.

AUTHOR CONTRIBUTIONS

Shayan Barootchi: conception and design of the study; analysis and interpretation of data; initial and final drafting of the work; final approval of the version to be published; accountable for all aspects of the work. Leonardo Mancini: initial and final drafting of the work; final approval of the version to be published; accountable for all aspects of the work. Teresa Heck: study design; illustrator of Figure 1; critical review of the manuscript; accountable for all aspects of the work. Martina Stefanini: manuscript preparation and the initial draft; final review of the work; accountable for all aspects of the work. Giovanni Zucchelli: initial design of the study; served as the gold standard examiner; gave final approval of the version to be published with critical manuscript review; accountable for all aspects of the work. Giulio Rasperini: manuscript preparation and the initial draft; final review of the work; accountable for all aspects of the work. Hom-Lay Wang: study design; critical review of the draft of the manuscript; final review; approval of the work and accountable for all aspects. Lorenzo Tavelli: study design; preparation and distribution of all cases; instruction to examiners; manuscript preparation; final review of the work; accountable for all aspects of the work.

DATA AVAILABILITY STATEMENT

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

CONFLICTS OF INTEREST

The authors do not have any financial interests, either directly or indirectly, in the products or information enclosed in the paper. The study was self-supported by the authors.

**ORCID**

Shayan Barootchi  <https://orcid.org/0000-0002-5347-6577>

Leonardo Mancini  <https://orcid.org/0000-0002-7030-155X>

Giulio Rasperini  <https://orcid.org/0000-0003-3836-147X>

Hom-Lay Wang  <https://orcid.org/0000-0003-4238-1799>

Lorenzo Tavelli  <https://orcid.org/0000-0003-4864-3964>

REFERENCES

- Elani HW, Starr JR, Da Silva JD, Gallucci GO. Trends in dental implant use in the U.S., 1999-2016, and projections to 2026. *J Dent Res*. 2018;97:1424-1430.
- Buser D, Sennerby L, De Bruyn H. Modern implant dentistry based on osseointegration: 50 years of progress, current trends and open questions. *Periodontol*. 2000;73:7-21.
- Barootchi S, Askar H, Ravida A, Gargallo-Albiol J, Travan S, Wang HL. Long-term clinical outcomes and cost-effectiveness of full-arch implant-supported zirconia-based and metal-acrylic fixed dental prostheses: a retrospective analysis. *Int J Oral Maxillofac Implants*. 2020;35:395-405.
- Barootchi S, Ravida A, Tavelli L, Wang HL. Nonsurgical treatment for peri-implant mucositis: a systematic review and meta-analysis. *Int J Oral Implantol (Berl)*. 2020;13:123-139.
- Wen SC, Barootchi S, Huang WX, Wang HL. Surgical reconstructive treatment for infraosseous peri-implantitis defects with a submerged healing approach: a prospective controlled study. *J Periodontol*. May 29. <https://doi.org/10.1002/JPER.21-0161>.
- Barootchi S, Tavelli L, Majzoub J, Chan HL, Wang HL, Kripfgans OD. Ultrasonographic tissue perfusion in peri-implant health and disease. *J Dent Res*. 2021;220345211035684.
- Barootchi S, Wang HL. Peri-implant diseases: current understanding and management. *Int J Oral Implantol (Berl)*. 2021;14:263-282.
- Wang II, Barootchi S, Tavelli L, Wang HL. The peri-implant phenotype and implant esthetic complications. Contemporary overview. *J Esthet Restor Dent*. 2021;33:212-223.
- Zucchelli G, Barootchi S, Tavelli L, Stefanini M, Rasperini G, Wang HL. Implant soft tissue dehiscence coverage esthetic score (IDES): a pilot within- and between-rater analysis of consistency in objective and subjective scores. *Clin Oral Implants Res*. 2021;32:349-358.
- Cosyn J, De Bruyn H, Cleymaet R. Soft tissue preservation and pink aesthetics around single immediate implant restorations: a 1-year prospective study. *Clin Implant Dent Relat Res*. 2013;15:847-857.
- Burkhardt R, Joss A, Lang NP. Soft tissue dehiscence coverage around endosseous implants: a prospective cohort study. *Clin Oral Implants Res*. 2008;19:451-457.
- Sanz-Martin I, Regidor E, Navarro J, Sanz-Sanchez I, Sanz M, Ortiz-Vigon A. Factors associated with the presence of peri-implant buccal soft tissue dehiscences: a case-control study. *J Periodontol*. 2020;91:1003-1010.
- Cairo F, Nieri M, Cavalcanti R, et al. Marginal soft tissue recession after lateral guided bone regeneration at implant site: a long-term study with at least 5 years of loading. *Clin Oral Implants Res*. 2020;31:1116-1124.
- Anderson LE, Inglehart MR, El-Kholy K, Eber R, Wang HL. Implant associated soft tissue defects in the anterior maxilla: a randomized control trial comparing subepithelial connective tissue graft and acellular dermal matrix allograft. *Implant Dent*. 2014;23:416-425.
- Mazzotti C, Stefanini M, Felice P, Bentivogli V, Mounssif I, Zucchelli G. Soft-tissue dehiscence coverage at peri-implant sites. *Periodontol*. 2000;77:256-272.
- Zucchelli G, Mazzotti C, Mounssif I, Mele M, Stefanini M, Montebugnoli L. A novel surgical-prosthetic approach for soft tissue dehiscence coverage around single implant. *Clin Oral Implants Res*. 2013;24:957-962.
- Frisch E, Ratka-Kruger P. A new technique for peri-implant recession treatment: partially epithelialized connective tissue grafts. Description of the technique and preliminary results of a case series. *Clin Implant Dent Relat Res*. 2020;22:403-408.
- Rocuzzo M, Gaudio L, Bunino M, Dalmasso P. Surgical treatment of buccal soft tissue recessions around single implants: 1-year results from a prospective pilot study. *Clin Oral Implants Res*. 2014;25:641-646.
- Romandini M, Pedrinaci I, Lima C, Soldini MC, Araoz A, Sanz M. Prevalence and risk/protective indicators of buccal soft tissue dehiscence around dental implants. *J Clin Periodontol*. 2021;48:455-463.
- Stefanini M, Marzadori M, Tavelli L, Bellone P, Zucchelli G. Peri-implant papillae reconstruction at an esthetically failing implant. *Int J Periodontics Restorative Dent*. 2020;40:213-222.
- Zucchelli G, Tavelli L, Stefanini M, et al. Classification of facial peri-implant soft tissue dehiscence/deficiencies at single implant sites in the esthetic zone. *J Periodontol*. 2019;90:1116-1124.
- O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med*. 2014;89:1245-1251.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
- Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation—A discussion and demonstration of basic features. *PLoS One*. 2019;14:e0219854.
- Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163.
- Signorell A. DescTools: Tools for Descriptive Statistics. 2019.
- Kuznetsova AB, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82:1-26.
- Eddelbuette DF. Rcpp: seamless R and C++ integration. *J Stat Softw*. 2011;40:1-18.
- Eddelbuette D. *Seamless R and C++ Integration with Rcpp*. Springer; 2013.
- Eddelbuette DB. Extending R with C++: a brief introduction to rcpp. *PeerJ Preprints*. 2017;5:e3188v3181.
- Bürkner PC. brms: an R package for Bayesian multilevel models using stan. *J Stat Softw*. 2017;80:1-28.
- Bürkner PC. Advanced Bayesian multilevel modeling with the r package brms. *R J*. 2018;10:395-411.
- Wickham HH. L. tidy: tidy messy data. 2020.
- Kay M. tidybayes: tidy data and geoms for Bayesian models. 2020.



35. Wickham H. ggplot2: elegant graphics for data analysis. 2016.
36. Cairo F, Nieri M, Cincinelli S, Mervelt J, Pagliaro U. The interproximal clinical attachment level to classify gingival recessions and predict root coverage outcomes: an explorative and reliability study. *J Clin Periodontol.* 2011;38:661-666.
37. Cairo F, Nieri M, Cattabriga M, et al. Root coverage esthetic score after treatment of gingival recession: an interrater agreement multicenter study. *J Periodontol.* 2010;81:1752-1758.
38. Wessels R, De Roose S, De Bruyckere T, et al. The Mucosal Scarring Index: reliability of a new composite index for assessing scarring following oral surgery. *Clin Oral Investig.* 2019;23:1209-1215.
39. Cortellini P, Bissada NF. Mucogingival conditions in the natural dentition: narrative review, case definitions, and diagnostic considerations. *J Periodontol.* 2018;89(Suppl 1):S204.
40. Cairo F, Rotundo R, Miller PD, Pini Prato GP. Root coverage esthetic score: a system to evaluate the esthetic outcome of the treatment of gingival recession through evaluation of clinical cases. *J Periodontol.* 2009;80:705-710.
41. Barootchi S, Tavelli L, Di Gianfilippo R, et al. Long term assessment of root coverage stability using connective tissue graft with or without an epithelial collar for gingival recession treatment. A 12-year follow-up from a randomized clinical trial. *J Clin Periodontol.* 2019;46:1124-1133.
42. Tavelli L, Barootchi S, Di Gianfilippo R, et al. Acellular dermal matrix and coronally advanced flap or tunnel technique in the treatment of multiple adjacent gingival recessions. A 12-year follow-up from a randomized clinical trial. *J Clin Periodontol.* 2019;46:937-948.
43. Barootchi S, Tavelli L, Gianfilippo RD, et al. Acellular dermal matrix for root coverage procedures: 9-year assessment of treated isolated gingival recessions and their adjacent untreated sites. *J Periodontol.* 2021;92:254-262.
44. Cairo F, Barootchi S, Tavelli L, et al. Aesthetic-And patient-related outcomes following root coverage procedures: a systematic review and network meta-analysis. *J Clin Periodontol.* 2020;47:1403-1415.
45. Barootchi S, Tavelli L, Zucchelli G, Giannobile WV, Wang HL. Gingival phenotype modification therapies on natural teeth: a network meta-analysis. *J Periodontol.* 2020;91:1386-1399.
46. Gil S, de la Rosa M, Mancini E, et al. Coronally advanced flap achieved higher esthetic outcomes without a connective tissue graft for the treatment of single gingival recessions: a 4-year randomized clinical trial. *Clin Oral Investig.* 2021;25:2727-2735.
47. Isaia F, Gyurko R, Roomian TC, Hawley CE. The root coverage esthetic score: intra-examiner reliability among dental students and dental faculty. *J Periodontol.* 2018;89:833-839.
48. Zucchelli G, Tavelli L, Stefanini M, Barootchi S, Wang HL. The coronally advanced flap technique revisited: treatment of peri-implant soft tissue dehiscences. *Int J Oral Implantol (Berl).* 2021;14:351-365.
49. Zucchelli G, Tavelli L, Barootchi S, Stefanini M, Wang HL, Cortellini P. Clinical remarks on the significance of tooth malposition and papillae dimension on the prediction of root coverage. *Int J Periodont Rest Dent.* 2020;40:795-803.

How to cite this article: Barootchi S, Mancini L, Heck T, et al. Reliability assessment of the classification for facial peri-implant soft tissue dehiscence/deficiencies (PSTDs): A multi-center inter-rater agreement study of different skill-level practitioners. *J Periodontol.* 2022;93:1173–1182. <https://doi.org/10.1002/JPER.21-0606>