The Prognostic Performance of the 2017 World Workshop Classification on Staging and Grading of Periodontitis compared to the British Society of Periodontology's implementation

Himabindu Dukka[1]; Thomas Dietrich[2,5]; Muhammad H. A. Saleh[1,3]; Giuseppe Troiano[4]; Zehra Yonel[2,5]; Andrea Ravidà[3]; Hom-Lay Wang[3], Henry Greenwell[1]; Iain Chapple[2,5]

1. Department of Periodontics, University of Louisville School of Dentistry, Louisville, KY, USA.

2.Department of Periodontology, The School of Dentistry, University of Birmingham, Birmingham, UK

3.Department of Periodontics and Oral Medicine, University of Michigan School of Dentistry, Ann Arbor, MI, USA.

4. Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy.

5.Birmingham Community Health Foundation NHS Trust, Birmingham, UK.

Corresponding author:
Professor Iain L C Chapple
Director of Research
Institute of Clinical Sciences,
College of Medical & Dental Sciences
The University of Birmingham
5 Mill Pool Way,
Edgbaston, Birmingham B5 7EG
United Kingdom
Tel: 0121 4665129
Sec: 0121 4665486
Email: I.L.C.Chapple@bham.ac.uk

Abstract word count: 264
Total word count: 3128
Total number of tables and figures: 2 figures, 3 tables.
Supplemental material: 1 table
Number of references: 33
Running title: Comparison of Stage and Grade to the BSP.

Author contributions:
Study conception and design: M.S., G.T.
***Analysis and interpretation of the data*:** H.D., M.S., G.T., T.D.
Data collection: M.S., H.D., A.R., M.Q.
Drafting of the manuscript: M.S., H.D., H.G.
Critical revision of the manuscript H.G., HL.W, T.D., I.C.
All authors gave their final approval and agreed to be accountable for all aspects of the work.

One sentence summary: The overall prognostic performance of the two classification systems was comparable with excellent predictive ability.

Conflict of Interest and Source of Funding: The authors do not have any financial interests, either directly or indirectly, in the products or information listed in the paper.

Abstract

Background: The British Society of Periodontology (BSP) implemented a simplified version of the 2017 World Workshop Classification (WWC) on Staging and Grading of periodontitis, for use in UK clinical practice. The aim of this study was to assess the long-term (>10 years) prognostic capability of BSP's implementation (BSP-i) compared to the 2017 WWC, using periodontal-related tooth loss (TLP) as a disease outcome.

Methods: Data on medical history, smoking status, and clinical periodontal parameters were retrieved from 270 patients who received non-surgical and surgical periodontal therapy from 1966 to 2007. Each patient received a baseline diagnosis according to the 2017 WWC and the BSP-i guidelines for implementation. Univariate multilevel Cox regression frailty models were performed to analyze the association between variables with TLP. A post-hoc comparison with Bonferroni correction was performed to analyze interclass comparisons. The prognostic performance of both systems was analyzed using Harrell's C index.

Results: The prognostic performance of both systems was very similar (0.922 for the 2017 WWC and 0.925 for the BSP-i). The singular prognostic performance of BSP Stage was slightly higher than that of 2017 WWC Stage (0.9212 vs 0.9188), while the 2017 WWC Grade showed a slightly better performance than BSP Grade (0.9175 vs 0.9155). BSP-i's Extent performed better than the 2017 WWC Extent (0.9203 vs 0.9098), however in the 2017 WWC Extent, the class "localized" was associated with a better prognosis than "generalized."

Conclusion: The overall prognostic performance of the two systems was excellent, with both systems having a Harrell's C index score of more than 0.92.

Keywords (MeSH): Attachment Loss, periodontitis, risk factor assessment, tooth loss, validation study

Introduction

Periodontitis is a plaque-induced multifactorial disease that is chronic in nature. It is initiated by the emergence of a dysbiosis within the dental plaque biofilm[1] and ultimately bone and attachment loss (AL) results from a disproportionate host immune-inflammatory response to the dysbiosis[2,3]. The host response is determined by genetic, epigenetic, lifestyle, environmental and behavioral risk factors[4] which makes risk prediction for disease progression challenging. Like many other chronic diseases, there is no cure for periodontitis and supportive periodontal therapy (SPT), also known as 'periodontal maintenance,' is paramount to prevent future deterioration[5]. Periodontal therapy should not only involve eliminating and/or controlling the associated symptoms but also include controlling the predisposing and modifying factors (local and systemic risk factors) that impact disease progression[6]. For this reason, patient risk assessment needs to be performed at multiple levels namely, the patient/systemic level, mouth level, tooth and site level[7].

The European Federation of Periodontology (EFP) and American Academy of Periodontology (AAP) jointly implemented the concept of risk assessment in the 2017 international classification system for periodontal diseases[8]. The new classification system employs a protocol for disease Staging and Grading, which has long been utilized in the diagnosis and treatment of malignant tumors, where the Stage is a measure of the size of the lesion (tumor) at the point of diagnosis and the Grade measures rate of cancer progression, based upon histological and/or molecular features of the lesion[9,10]. The 2017 classification also includes an in-built prognostication system, reinforcing the significance of risk assessment in comprehensive patient evaluation, and provides the necessary framework for inclusion of biomarker-based diagnostics for enhanced prognostication and risk stratification when future validated biomarker panels become available[11]. This prognostic capability of Staging and Grading was validated in a recent study, associating increased class severity (higher Stage or Grade) with increased tooth loss due to periodontitis (TLP)[12].

In the United Kingdom, the British Society of Periodontology (BSP) also convened an implementation group to develop guidance on how the new classification system could be simply implemented in clinical practice within the public 'national health service' (NHS). The BSP adopted a reductionist approach, as they felt the proposed system needed to be simplified if it was to be adopted in the UK's general practice environment[13]. Additionally, for the classification to be accepted within the NHS system, it needed to integrate established screening tools in UK like the 'Basic Periodontal Examination' (BPE) and clinical periodontal parameters like probing pocket depth (PPD) and bleeding on probing (BOP). The BSP identified several challenges in implementing the 2017 classification in general dental practice and hence made minimal but key adaptations to the classification as described by Tonetti et al.[11,13]. Staging was implemented using a singular factor, radiographic bone loss for determining severity (or AL where contemporary radiographs are not available), excluding all other complexity factors, and changing the Stage IV threshold to 'bone loss within the apical 1/3rd of the root. For Grade, changes were made to Grade B threshold, and systemic complexity factors were not considered within the grading approach but documented separately noting them as risk factors as a part of the "diagnostic statement"[13] (Table 1).

The BSP-i was rapidly integrated into national NHS policy and protocols[14,15], however to the best of our knowledge, there are currently no published studies that have evaluated the reliability of BSP-version of the 2017 classification of periodontitis. Hence, the aim of this study was to assess the long-term (>10 years) prognostic capability of the BSP's implementation compared to the original 2017 World Workshop Classification (WWC) on Staging and Grading of periodontitis, using TLP as a definite outcome.

Materials and Methods

This study was conducted in agreement with the 1975 Declaration of Helsinki (World Medical Association [WMA], 1975) as most recently revised in 2013 (WMA, 2013)[16]. The study was approved by the University of Michigan Medical School Institutional Review Board (IRBMED) with study identifier HUM00157260/HUM00160933. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were followed during the preparation of the manuscript[17].

Study population

This study was conducted on a periodontitis patient population who received non-surgical and surgical periodontal therapy from January 1966 to January 2007 at the University of Michigan School of Dentistry. Physical and digital records of patients were screened and evaluated by three examiners (MQ, AR, and MS). The following eligibility criteria were established:

- Patients met the case definition of periodontitis as defined by Tonetti et al.[11]
- Patients had at least one session of scaling and root planing (SRP)/diseased area with or without additional surgery if needed) and maintained for ≥10 years after active therapy at the University of Michigan School of Dentistry.
- Complete patient charts with full mouth radiographic series of diagnostic quality (taken within ≤12 months from the baseline/initial periodontal examination).
- Complete medical history recorded at baseline periodontal examination.
- Patients received at least one visit of supportive periodontal therapy (SPT)/year throughout the entire follow-up period.
- Patients whose teeth have been extracted at the University of Michigan School of Dentistry.

Patients who did not meet the *a priori* set criteria were excluded from the study. Demographic patient information, periodontal status, total number of SPT visits/year and relevant medical history (smoking and diabetes mellitus) were collected. Tooth loss (TL) data was analyzed in two stages. First, the overall tooth loss (OTL) was calculated by deducting the number of natural teeth present at the last follow-up visit (T1) from the number present at baseline periodontal examination (T0). Second, the cause of extraction of each tooth was identified as determined by patients' charts and calculating the time the tooth stayed in function until extraction. The reason of tooth loss should have been stated in patients' notes at the time of extraction, and charts that did mention the cause of tooth loss were excluded from the analysis. Only teeth extracted due to periodontal reasons (TLP) were considered in the current analysis. Teeth that were extracted due to reasons other than TLP were censored in the survival analysis.

Percentage of radiographic bone loss (RBL, in %) was primarily measured from periapical radiographs[18]. Probing pocket depths (PPD) and AL were evaluated at six sites per tooth. Information about tooth mobility, drifting or flaring (tooth migration), bite collapse, parafunctional habits, chewing difficulties, masticatory dysfunction, and plaque accumulation were collected from patient records when available. The number of teeth lost that was attributable to periodontitis as defined by Sanz et al (0, ≤4 or ≥5) were reported[19].

Data collection and patient classification

Before classifying patients, the case definition for periodontitis as defined by the 2017 WWC[11] was confirmed. Then, each patient received a baseline diagnosis according to either the 2017 WWC [8] or BSP guidelines for implementation of the classification: for the 2017 WWC system, Stage: I, II, III, or IV; Grade A, B, or C were assigned to all patients by a single investigator (MS), after being calibrated by an expert author (HG). All teeth lost during active periodontal treatment such as teeth that were deemed hopeless at the patient screening were not considered when Staging the patients and thus excluded from the study. The classification of Staging and Grading according to the BSP system was performed by two investigators (HD), (MS). Table 1 shows a direct comparison of the parameters used in both classifications.

The 2017 WWC has clarified in a more recent publication that the extent of periodontitis involvement is to be considered localized if ≤30% of teeth are found to be at the Stage-defining severity level. Whereas the Extent would be generalized if >30% of teeth were at the Stage-defining severity level[19]. All Extent allocation in the present study was made based on this clarification.

Statistical Analysis

The following patient-level variables were included in the analysis: patient demographics, age, gender; Stage, Grade and Extent according to the 2017 WWC system; Stage, Grade and Extent according to the BSP implementation system. In addition, the following tooth-level variables were included: time occurring from baseline to the last follow-up, status at the last follow-up (Teeth present, Teeth lost to periodontitis, Teeth lost for reasons other than periodontitis). In addition, the number of SPT sessions undertaken by the patient during follow-up was extracted from records and included as a variable in this study. Univariate multilevel Cox regression frailty models were performed in order to analyze the association of variables with periodontitis-related tooth loss. Hence, the influence of confounding factors on the prognostic performance of variables from the classification systems, multilevel multivariable Cox regression frailty models (including Age, Gender and number of maintenance sessions undertaken by the patients during follow-up as confounding factors), addressing the clustering of teeth within subjects, were built. A post-hoc comparison with Bonferroni test was performed to analyze interclass comparison after the Cox regression. The prognostic performance of both BSP and 2017 WWC variables was analyzed by calculating Harrell's C index, Akaike and Bayesian Information Criterion (AIC and BIC) from the multivariable multilevel models. An internal validation of the prognostic measurements calculated in the univariate analysis was also undertaken by means of k-fold cross-validation (crossfold command in STATA).

Results

Difference in the parameter used in both classification systems

The difference in parameters used for determining either the Stage or Grade of Periodontitis between the 2017 WWC and BSP-i are all shown in Table 1.

Demographic data of the cohort and allocation according to the two classification systems

Sufficient information for a total of 270 patients fulfilled the eligibility criteria and as included in this analysis, making up a total of 6833 teeth. The average age in the cohort was 42.9 ± 11.8 years, of these 143 (53.0%) were male and 127 (47.0%) were female. The total number of teeth lost due to periodontal disease (TLP) over the follow-up period was 318 teeth (4.6%). Patients were allocated into categories according to both 2017 WWC and BSP-i (Figure 1). The BSP-i allocated significantly more patients in Stage II (42.2%) and also had more Grade B patients (66.3%). The most significant differences were found for Extent, with BSP-i system classifying more patients (43.3%) as generalized.

The post-hoc inter-rater reliability for the BSP-i system was measured using kappa coefficient (K). The K-agreement result for the Stage was 0.79; Grade: 1.0; and Extent: 0.57. Differences in allocation between the two systems were evident for both Stage, Grade and Extent. In particular, the BSP-i significantly allocated more patients in Stage II (114/220, 42.2%) and less in Stage III (92/220, 34.0%) compared to the 2017 WWC (28.2% and 50.3%). Difference in allocation between BSP-i and 2017 WWC were also evident for Grade A (10.7% vs 17.8%) and Grade B (66.3% vs 54.4%). Focusing on the Extent of periodontitis, the BSP-i system classified more patients as generalized (43.3%) compared to the WWC2017 system (28.9%).

Prognostic analysis of variables according to BSP-i and WWC 2017 systems

The univariate survival analysis categories within 2017 WWC Stage and Grade were significantly correlated with TLP, while the Extent did not (Table 2). Kaplan-Meier figures graphically showed the prognostic stratification at the univariate analysis (Figure 2). After adjusting for the confounding effects of age, gender and number of periodontal maintenance sessions, 2017 WWC Stage III and IV were correlated with a worse prognosis, a similar trend was observed for 2017 WWC Grade C ($p = 0.070$). 2017 WWC Extent did not correlate with TLP even in the multivariate analysis. Focusing on the BSP-i system, all the analyzed variables (Stage, Grade and Extent) were significantly correlated with TLP both at univariate and multivariate analysis.

A direct comparison of the two systems was performed by analyzing the prognostic performance of the corresponding variables in the multivariate models. The general prognostic performance (including all variables together in the model) of the two systems was very similar (0.922 for the 2017 WWC and 0.925 for BSP-i). Table 3 shows the impact of controlling each of the confounders on the prognostic performance. The singular prognostic performance of the BSP-i Stage was slightly higher than that of 2017 WWC Stage (0.9212 vs 0.9188), while the 2017 WWC Grade showed a slightly better performance than the BSP-i Grade (0.9175 vs 0.9155). Focusing on the extent of the disease the BSP-i Extent performed better than the 2017 WWC Extent (0.9203 vs 0.9098), however it is important to note that in the 2017 WWC Extent, the class “localized” was associated with a better prognosis than “generalized.”

Discussion

Our results revealed that the different risk categories of periodontitis as defined by both the 2017 WWC and the BSP-i were associated with different risk classes. Basically, the higher the Stage or Grade is, the greater the risk is for TLP. However, with small exceptions, different class severities of Stage and Grade had a stronger correlation with TLP in the BSP-i when compared to the original 2017 WWC (Table 2). Both models showed excellent overall prognostic performance (0.922 for the 2017 WWC and 0.925 for the BSP-i). Although the overall prognostic performance was very similar, the prognostic performance of the BSP-i Stage and Extent was slightly better than the 2017 WWC Stage, while the 2017 WWC Grade demonstrated a better performance than the BSP-i Grade.

Patient allocation between the two systems was slightly different (Figure 1). The most significant differences were found for Extent. The BSP-i system classified 43.3% of the patients as generalized compared to the 2017 WWC system 28.9%. The approach by which the Extent component of the 2017 WWC had to be used was clarified in a recent publication[19], where the number of teeth at the Stage-defining severity were considered with a cut-off percentage of 30% for a localized versus a generalized extent[19]. In the current analysis, the majority of Stage I or II cases (mild and moderate periodontitis) exhibited a generalized extent of disease. Interestingly, the 2017 WWC Extent did not correlate with TLP in either the univariate or the multivariate analysis. Another important difference was that the 2017 WWC allocated more cases to Stage III. This may be due to the description of the complexity factors (vertical BL ≥3 mm; furcation II or III; PD ≥6 mm; or moderate ridge defects) in the 2017 WWC, which may have driven more readily the allocation to Stage III versus a singular factor of radiographic evaluation in the BSP-i.

It is worthy of note that the risk factors included in the 2017 WWC such as smoking and diabetes mellitus while in the BSP-I, both are considered as the separate risk entities in addition to the Stage, Grade and Extent, and not as a part of the classification process itself. The BSP-i system assumes that the percentage of radiographic bone loss/age ratio captures the historical disease susceptibility, with all patients’ risk factors and indicators leading to it, such as smoking and poorly controlled diabetes. This, however, does not take into consideration future disease susceptibility. Nevertheless, disease susceptibility may change if a patient’s smoking habit changes (from heavy to light and vice versa)[21]. The same could be expected for patients with diabetes[22].

The results of the multilevel-multivariate analysis demonstrated the number of SPT sessions attended by patients represented the best predictor for tooth survival. This agrees with the overwhelming body of evidence available[23]. Irregular compliance with SPT has been consistently associated with an increased risk of tooth loss [25-26]. Additionally, SPT frequency and patient compliance seemed to diminish the detrimental effects of residual PPD and smoking, on maintenance therapy outcomes in terms of tooth loss[27,28]. This is demonstrated by the low rate of TLP (4.6%) encountered in the current study.

The present analysis has several strengths. Instead of using overall tooth loss (OTL) that is frequently employed in similar studies, only TLP was used. The key parameters that are usually used to gauge periodontitis severity (similar to those used for Stage and Grade) were not found to affect OTL[20]. A recent long-term

investigation has demonstrated that the Staging and Grading systems are indeed prognostic for TLP, but not OTL[12]. Moreover, the current study only assessed teeth that were lost during SPT and did not include any teeth that were extracted during the cause related phase of therapy.

One of the limitations of the present study was that the authors were unable to perform a classical external validation study of the two classification systems by calculating discrimination and calibration, as no prior model has been developed and the pre-requisite model parameters are not available to test[30-32]. Therefore, it was decided that a formal validation study would not be possible. The authors opted to evaluate the single prognostic performance of each variable in a univariate analysis and multivariate analysis was performed with the aim to adjust the results of the univariate analysis for confounding factors. The overall analysis was performed to evaluate the statistical significance of variables included and evaluate if they were independently associated with a higher risk of tooth loss, to allow comparison of similar parameters in the two classifications and analyze how they performed (e.g., Stage in the BSP-i vs Stage in the WWC). The authors recognize the limitations associated with this method and, to reduce the possibility of an over-optimistic performance, internal data validation techniques were applied by means of k-fold cross-validation (Supplementary table 1). Interestingly, the results of the internal validation confirmed the previously obtained findings, excluding the possibility of an overoptimistic performance of the initial models.

In addition, other limitations are related to the absence of an *a priori* calculation of the sample size and to the possibility of selection bias in the study cohort. As this study was not carried out using a prospective cohort of participants, the available sample size was pre-determined. However, our cohort respected the sample size requirements outlined by Collins et al for the external validation of time-to-event data[33]. In particular, a total of 318 TLs events were present in our cohort, exceeding the ideal threshold of 200 events required to provide an adequate statistical power[33].

The current study was conducted using dental school data records. A range of operators treated those patients, including undergraduate and graduate dental students, and their instructors. All those bring their own biases. This could lead to different criteria being applied clinically for the need for extractions, or how a "periodontally hopeless" tooth is defined. The inclusion of data from many years could also have led to some systematic bias caused by changing perspectives on the possibility of rescuing a tooth as compared to implant placement. This study set the limit of ≥10 years of regular SPT as a criterion for inclusion. This selectively excluded patients who either died or lost their entire dentition due to rapid periodontitis progression before they hit the 10-year mark. The ≥10 years SPT threshold ensured that an effect from TLP could be demonstrated, given the slow pattern of periodontitis progression[29].

The ideal system for risk assessment in everyday practice should be quick, simple, reliable and easy to understand for both the professional and the patient. The main rationale behind abridging the 2017 WWC by the BSP-i was to simplify the classification process for clinicians in practice, especially general practitioners and dental students. This could make treatment outcomes more predictable and improve our ability to share findings with our patients. A few examples of cases that demonstrate the practical implementation of the BPS-I were published recently.[34,35] More importantly, the results presented in this manuscript demonstrated that the BSP's pragmatic approach through implementing a reductionist model of the original 2017 WWC neither affected the class allocation nor the prognostic performance of the system. This should be considered as the most significant finding in terms of practicality of implementing this classification in general dental practice. Indeed, in July 2021 the NHS Business Services Authority embedded the BSP-I and associated stages of S3-Level guidelines for treatment within NHS statute for dental practices[36].

Conclusions

Different risk categories of periodontitis as defined by both the 2017 WWC and the BSP implementation of the classification were associated with different risk classes. The overall prognostic performance of the two systems was excellent, with both systems having a Harrell's C index score of more than 0.92.

Acknowledgment: Musa Qazi, Department of Periodontics and Oral Medicine, University of Michigan School of Dentistry, Ann Arbor, MI, USA, for assistance with data collection.

References

1. Hajishengallis G, Lamont RL. Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol Oral Microbiol*. 2012;27:409-419.

2. Bartold PM, Van Dyke TE. An appraisal of the role of specific bacteria in the initial pathogenesis of periodontitis. *J Clin Periodontol.* 2019;46:6-11.

3. Meyle J, Chapple ILC. Molecular aspects of the pathogenesis of periodontitis. *Periodontology 2000*. 2015;69:7–17.

4. Tonetti MS, Eickholz P, Loos BG, et al. Principles in prevention of periodontal diseases Consensus report of group 1 of the 11th European Workshop on Periodontology on effective prevention of periodontal and peri-implant diseases. *J Clin Periodontol.* 2015;42(Suppl 16):S5–S11.
5. Becker W, Becker BE, Berg LE. Periodontal treatment without maintenance. A retrospective study in 44 patients. *J Periodontol*. 1984;55:505-509.

6**.** Chapple ILC. Risk assessment in periodontal care: the principles. In: Chapple ILC, Papapanou P, eds. *Risk assessment in oral health. A Concise Guide for Clinical Application*. Switzerland: Springer Nature; 2020:77-88.

7. Salvi GE, Lawrence HP, Offenbacher S, Beck JD. Influence of risk factors on the pathogenesis of periodontitis. *Periodontology 2000*.1997;14:173-201.

8. Papapanou PN, Sanz M, Buduneli N, et al. Periodontitis: consensus report of workgroup 2 of the 2017 World Workshop on the classification of periodontal and peri-implant diseases and conditions. *J Clin Periodontol.* 2018;45(Suppl 20):S162-S170.

9. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67:93-99.

10. Orucevic A, Bell JL, King M, McNabb AP, Heidel RE. Nomogram update based on tailorx clinical trial results - oncotype dx breast cancer recurrence score can be predicted using clinicopathologic data. *Breast J*. 2019;46:116-125.

11. Tonetti MS, Greenwell H, Kornman KS. Staging and grading of periodontitis: framework and proposal of a new classification and case definition. *J Clin Periodontol.* 2018;45(Suppl 20):S149-S161.

12. Ravida A, Qazi M, Troiano G, et al. Using periodontal staging and grading system as a prognostic factor for future tooth loss: a long-term retrospective study. *J Periodontol.* 2020; 91:454-461.

13. Dietrich T, Ower P, Tank M, et al. Periodontal diagnosis in the context of the 2017 classification system of periodontal diseases and conditions: implementation in clinical practice. *Br Dent J*. 2019;226:16–22.

14. Commissioning Standard for Restorative Dentistry. NHS England & NHS Improvement. https://www.england.nhs.uk/wp-content/uploads/2019/07/commissioning-standard-for-restorative-dentistry-v1.pdf

15. Dental standard operating procedure: Transition to recovery. NHS England & NHS Improvement. https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/06/C1029-SOP-Transition-to-recovery-A-phased-transition-for-dental-practices-towards-the-resumption-of-the-full-r.pdf

16. World Medical Association. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191–2194.

17. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP; STROBE initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology*. 2007;18:800-804.

18. Pepelassi EA, Tsiklakis K, Diamanti-Kipioti A. Radiographic detection and assessment of the periodontal endosseous defects. *J Clin Periodontol.* 2000;27:224-230.

19. Sanz M, Papapanou PN, Tonetti MS, Greenwell H, Kornman K. Guest Editorial: clarifications on the use of the new classification of periodontitis. *J Clin Periodontol*. 2020;47:658-659.

20. Al-Shammari KF, Al-Khabbaz AK, Al-Ansari JM, Neiva R, Wang HL. Risk indicators for tooth loss due to periodontal disease. *J Periodontol.* 2005;76:1910-1918.

21. Ravida A, Troiano G, Qazi M, et al. Dose-dependent effect of smoking and smoking cessation on periodontitis-related tooth loss during 10 - 47 years periodontal maintenance: a retrospective study in compliant cohort. *J Clin Periodontol*. 2020;47:1132-1143.

22. Ramseier CA, Woelber JP, Kitzmann J, Detzen L, Carra MC, Bouchard P. Impact of risk factor control interventions for smoking cessation and promotion of healthy lifestyles in patients with periodontitis: a systematic review. *J Clin Periodontol*. 2020;47(Suppl 22):S90-S106.

23. Trombelli L, Simonelli A, Franceschetti G, Maietti E, Farina R. What periodontal recall interval is supported by evidence? *Periodontology 2000*. 2020;84:124-133.

24. Tada, S, Allen, PF, Ikebe, K, Matsuda, K, Maeda, Y. Impact of periodontal maintenance on tooth survival in patients with removable partial dentures. *J Clin Periodontol*. 2015;42:46–53.

25. Costa FO, Lages EJ, Cota LO, Lorentz TC, Soares RV, Cortelli JR. Tooth loss in individuals under periodontal maintenance therapy: 5-year prospective study. *J Periodontal Res.* 2014;49:121–128.

26. Salvi GE, Mischler DC, Schmidlin K, et al. Risk factors associated with the longevity of multi-rooted teeth. Long-term outcomes after active and supportive periodontal therapy. *J Clin Periodontol*. 2014;41:701–707.

27. Ramseier CA, Nydegger M, Walter C, et al. Time between recall visits and residual probing depths predict long-term stability in patients enrolled in supportive periodontal therapy. *J Clin Periodontol*. 2019;46:218-230.

28. Matuliene G, Pjetursson BE, Salvi GE, et al. Influence of residual pockets on progression of periodontitis and tooth loss: results after 11 years of maintenance. *J Clin Periodontol*. 2008;35:685-695.

29. American Academy of Periodontology task force report on the update to the 1999 classification of periodontal diseases and conditions. *J Periodontol*. 2015;86:835–838.

30. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol*. 2013;13:33.

31. Collins GS, DeGroot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14:40.

32. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol*. 2015;68:279-289.

33. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med.* 2016;35:214-26.

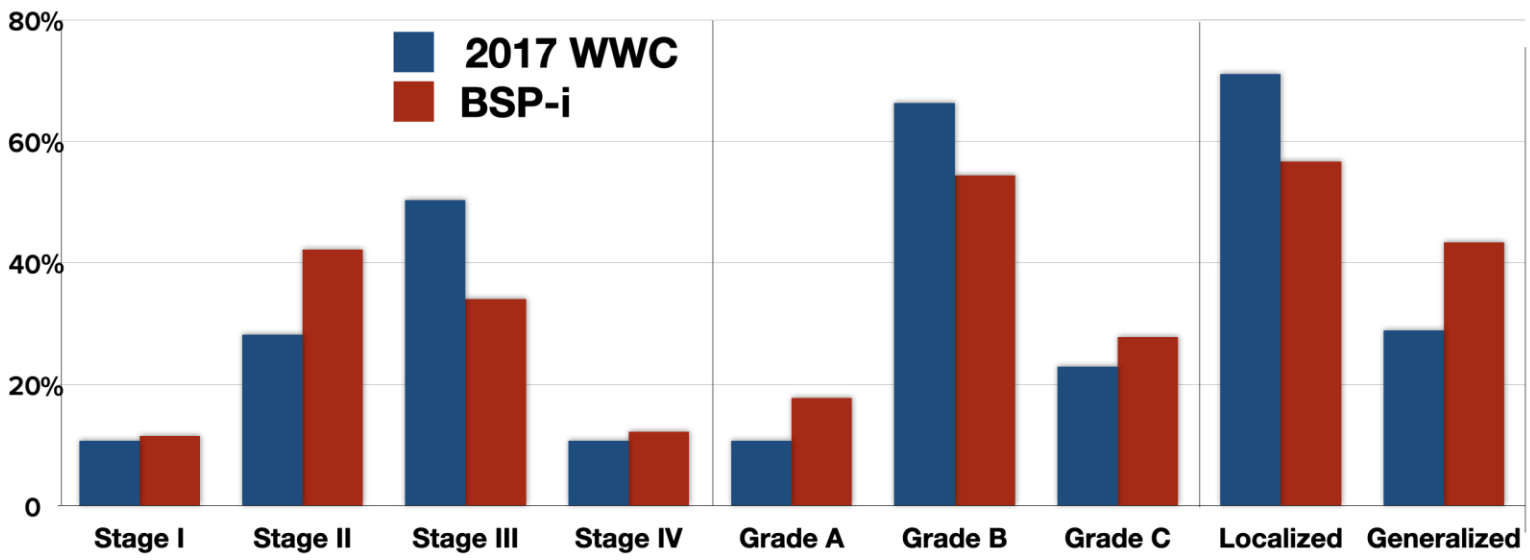
34. Walter C, Chapple ILC, Ower P, et al. Periodontal diagnosis in the context of the BSP implementation plan for the 2017 classification system of periodontal diseases and conditions: presentation of a patient with severe periodontitis following successful periodontal therapy and supportive periodontal treatment. *Br Dent J*. 2019;226:411-413.

35. Walter C, Ower P, Tank M, et al. Periodontal diagnosis in the context of the 2017 classification system of periodontal diseases and conditions: Presentation of a middle-aged patient with localized periodontitis. *Br Dent J*. 2019;226:98-100.

36. NHS. Avoidance of Doubt: provision of phased treatments. https://www.england.nhs.uk/wp-content/uploads/2018/02/B0615-Update-to-avoidance-of-doubt-provision-of-phased-treatments-300621-.pdf.

Figure Legends

Figure1: A frequency analysis portraying the frequency of occurrence of each class of Stage, Grade, or Extent in both the 2017 WWC and the BSP-i.



	Stage				Grade			Extent	
	I	II	III	IV	A	B	C	Loc	Gen
2017 WWC	29 (10.74%)	76 (28.15%)	136 (50.37%)	29 (10.74%)	29 (10.74%)	179 (66.30%)	62 (22.96%)	192 (71.11%)	78 (28.89%)
BSP-i	31 (11.48%)	114 (42.22%)	92 (34.07%)	33 (12.22%)	48 (17.78%)	147 (54.44%)	75 (27.78%)	117 (56.67%)	153 (43.33%)
p-value	0.839	*0.002**	*0.0006**	0.627	*0.035**	*0.011**	0.246	*0.0016**	

Fig. 1

Figure 2: Survival curves built for periodontal-related tooth loss (TLP on multilevel multivariate Cox Regression analysis adjusting for confounding factors such as: Age, Gender and number of maintenance visits.

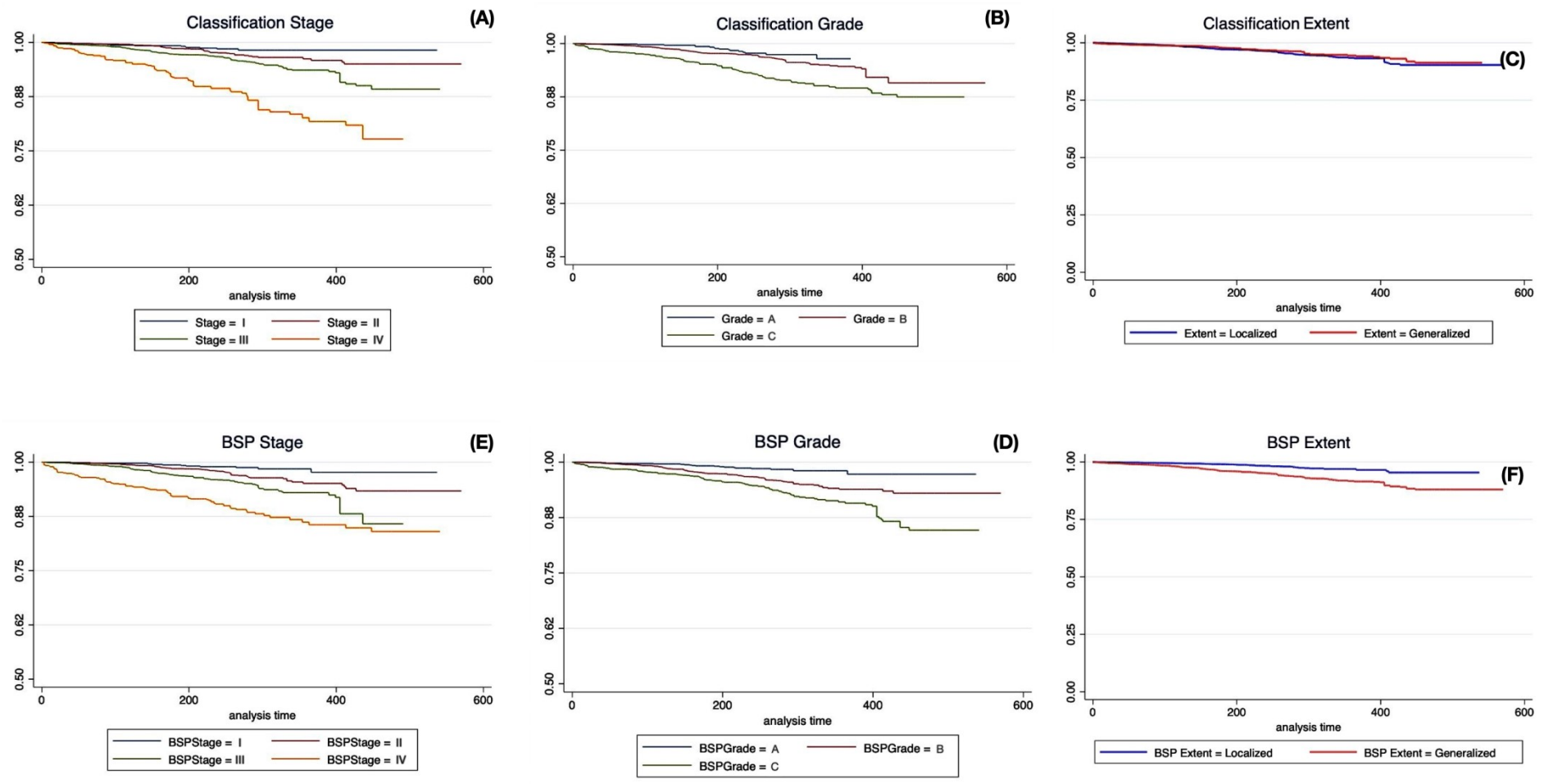
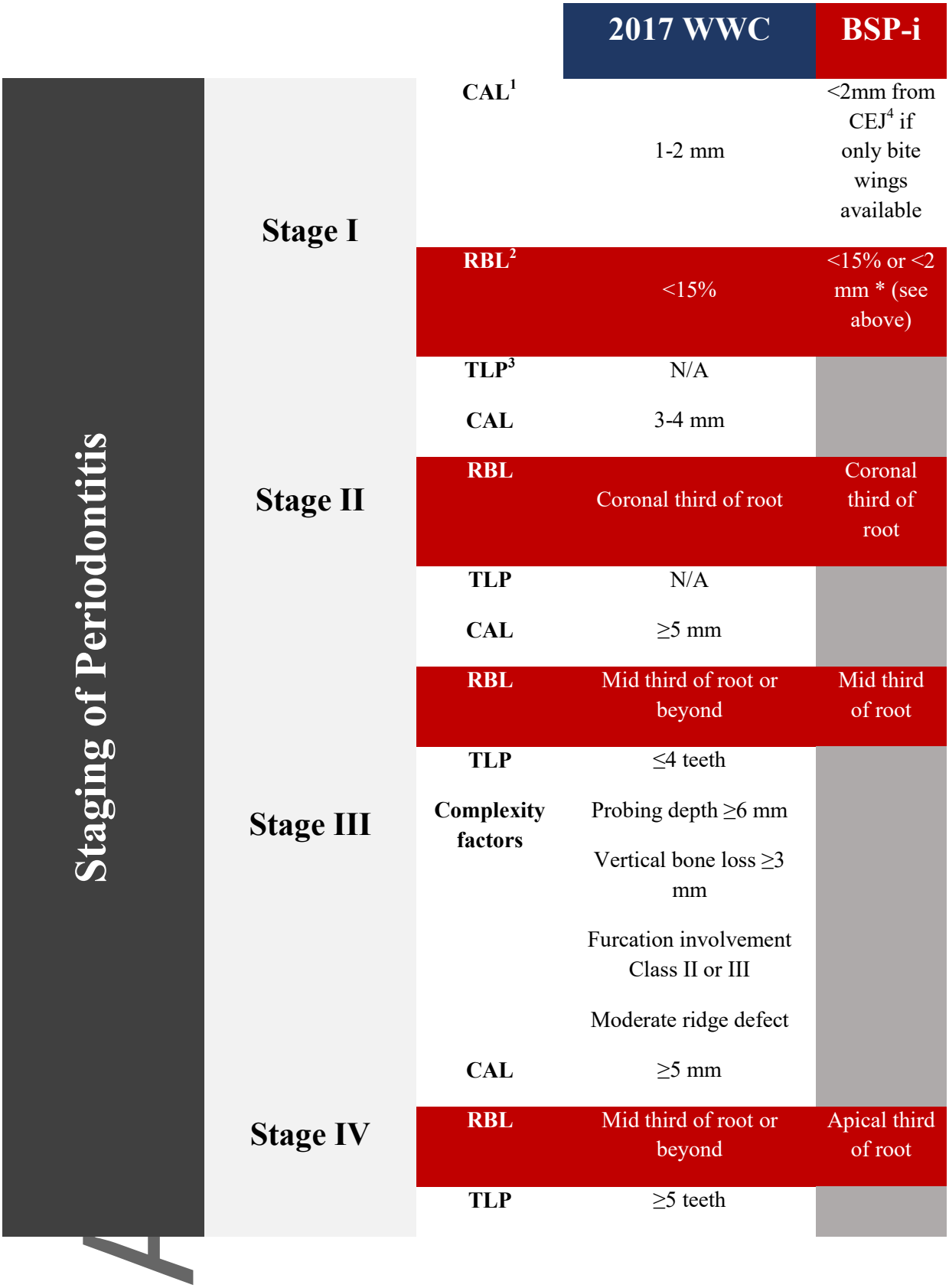


Fig. 2

Table 1: Comparison of the parameters used in the Staging and Grading of Periodontitis by 2017 WWC (The 2017 World Workshop classification) and BSP-i (The British Society of Periodontology implementation of the 2017 World Workshop classification).



Extent of Periodontitis		**Complexity factors**	Need for complex rehabilitation due to: Masticatory dysfunction Secondary occlusal trauma (tooth mobility degree ≥2) Severe ridge defect Bite collapse, drifting, flaring Less than 20 remaining teeth (10 opposing pairs)	
	Localized	≤30% of teeth are at the stage-defining severity level		≤30% of teeth involved
	Generalized	>30% of teeth are at the stage-defining severity level		>30% of teeth involved
Grading of Periodontitis		**Longitudinal CAL over 5 years**	Evidence of no loss	
		RBL/age	<0.25	<0.5
	Grade A	**Case phenotype**	Heavy biofilm deposits with low levels of destruction	
		Smoking	Non-Smoker	
		Diabetes	Normoglycemic / no diagnosis of diabetes	
		CRP[5]	<1 mg/L	
		Longitudinal CAL over 5 years	<2 mm	
	Grade B	**RBL/age**	0.25 to 1.0	0.5 to 1.0
		Case phenotype	Destruction commensurate with	

		biofilm deposits	
	Smoking	Smoker < 10 Cigarettes/day	
	Diabetes	HbA1c <7.0% in patients with diabetes	
	CRP	1 to 3 mg/L	
	Longitudinal CAL over 5 years	≥2 mm	
	RBL/age	>1.0	>1.0
Grade C	**Case phenotype**	Destruction exceeds expectation given biofilm deposits	
	Smoking	Smoker ≥10 cigarettes/day	
	Diabetes	HbA1c ≥7.0% in patients with diabetes	
	CRP	>3 mg/L	

*1= Clinical attachment loss; 2= Radiographic bone loss; 3= Periodontal-related tooth loss; 4= Cementoenamel junction; 5= C-reactive protein

Table 2: Multilevel univariate and multivariate Cox Regression frailty models for the two classification systems. Multivariate models were built including variables of the classification systems in conjunction with other confounding factors (Age, Gender and number of maintenance sessions).

Variables		Multilevel Univariate Analysis		Multilevel Multivariate Analysis	
		HR 95%(CI)	*p-value*	HR 95%(CI)	*p-value*
2017 WWC	I (ref)	1.00	-	1.00	-
Stage	II	1.81 (0.74-4.44)	0.196	2.43 (0.66-9.01)	0.183
	III	3.41 (1.47-7.89)	*0.004**	7.77 (2.28-26.54)	*0.001**
	IV	9.41 (3.53-25.06)	*0.000**	10.47 (2.58-42.5)	*0.001**
2017 WWC	A (ref)	1.00	-	1.00	-
Grade	B	1.90 (0.84-4.29)	0.120	1.24 (2.36-6.45)	0.723
	C	117.46 (29.6-462)	*0.001**	3.32 (0.90-12.2)	0.070
2017 WWC	1+ (ref)	1.00	-	1.00	-
Extent	2++	0.70 (0.50-0.99)	0.045*	0.71 (0.46-1.09)	0.120
BSP-i	I (ref)	1.00	-	1.00	-

Stage	II	2.88 (1.20-6.91)	*0.018**	3.94 (1.07-14.48) 0.039*
	III	5.56 (2.32-13.36)	*0.000**	8.33 (2.25-30.78) *0.001**
	IV	11.17 (4.21-29.67)	*0.000**	16.58 (4.06-67.66) *0.000**
BSP-i	A (ref)	1.00	-	1.00 -
Grade	B	3.35 (1.64-6.87)	*0.001**	4.84 (1.75-13.37) *0.002**
	C	4.79 (2.22-10.33)	*0.000**	9.49 (3.14-28.68) *0.000**
BSP-i	1 (ref)	1.00	-	1.00
Extent	2	3.65 (2.29-5.80)	0.000*	3.84 (1.91-7.69) *0.000**

* *Statistically significant*

[+]Localized; [++]Generalized

Table 3: Comparison of model prognostic stratification performance for TLP using measurements of model fit (Akaike's information criterion and Bayesian information criterion); and prognostic performance (Harrell's C-index). Higher Harrell's C-index and lower AIC and BIC, indicate better prognostic performance. Also included are values which are controlled for the confounding effect of age, gender and number of periodontal maintenance sessions.

Predictor	Prognostic performance from Multilevel Univariate Cox Regression Frailty models		
	Harrell's c-index	Akaike's Information Criterion (AIC)	Bayesian Information Criterion (BIC)
Age	0.567	5060	5067
Gender	0.513	5062	5067
2017 WWC_Stage	0.640	5039	5059
2017 WWC_Grade	0.608	5050	5064
2017 WWC _Extent	0.520	5059	5066
BSP-i_Stage	0.659	5037	5057
BSP-i_Grade	0.597	5050	5063
BSP-i_Extent	0.617	5035	5042
Maintenance Sessions	0.910	3805	3812
2017 WWC (Stage + Grade + Extent) *	0.680	5030	5070
BSP-i (Stage + Grade + Extent) *	0.691	5027	5068

*Multivariate Analysis

Supplementary table 1: Internal validation data techniques applied using k-fold cross-validation.