



Cite this: DOI: 10.1039/d3an01438j

Spectroscopic insights into peri-implant mucositis and peri-implantitis: unveiling peri-implant crevicular fluid profiles using surface enhanced Raman scattering†

Stefano Fornasaro,^a Antonio Rapani,^b Federica Farina,^b Marigona Ibishi,^b Giulia Pissoli,^b Claudio Stacchi,^b Valter Sergo,^c Alois Bonifacio,^c Roberto Di Lenarda^b and Federico Berton^b

The precise identification and differentiation of peri-implant diseases, without the need for intrusive procedures, is crucial for the successful clinical treatment and overall durability of dental implants. This work introduces a novel approach that combines surface-enhanced Raman scattering (SERS) spectroscopy with advanced chemometrics to analyse peri-implant crevicular fluid (PICF) samples. The primary purpose is to offer an unbiased evaluation of implant health. A detailed investigation was performed on PICF samples obtained from a cohort of patients exhibiting different levels of peri-implant health, including those with healthy implants, implants impacted by peri-implantitis, and implants with peri-implant mucositis. The obtained SERS spectra were analysed using canonical-powered partial least squares (CPPLS) to identify unique chemical characteristics associated with each inflammatory state. Significantly, our research findings unveil the presence of a common inflammatory SERS spectral pattern in cases of peri-implantitis and peri-implant mucositis. Furthermore, the SERS-based scores obtained from CPPLS were combined with established clinical scores and subjected to a linear discriminant analysis (LDA) classifier. Repeated double cross-validation was used to validate the method's capacity to discriminate different implant conditions. The integrated approach showcased high sensitivity and specificity and an overall balanced accuracy of 92%, demonstrating its potential to serve as a non-invasive diagnostic tool for real-time implant monitoring and early detection of inflammatory conditions.

Received 21st August 2023,
Accepted 26th December 2023

DOI: 10.1039/d3an01438j

rsc.li/analyst

Introduction

Dental implants, biocompatible metal devices surgically inserted into the residual bone crest of edentulous sites, have demonstrated favorable long-term survival rates in supporting dental prostheses. However, increasing clinical observations of mechanical and biological complications associated with functional implants have raised concerns over their long-term success.¹ Mechanical complications arising from occlusal forces, can lead to issues such as prosthetic screw loosening, prosthetic screw fracture, or fixture fracture. Biological compli-

cations, particularly biofilm-associated peri-implant diseases, including peri-implant mucositis (PIM) and peri-implantitis (PI), have also become prevalent.² PIM is characterized by soft tissue inflammation around the implant without loss of supporting bone,^{3,4} while PI involves inflammation in the peri-implant connective tissue accompanied by progressive bone loss.⁵ Studies estimate the prevalence of PIM and PI to be around 43% and 22%, respectively, based on meta-analysis and comprehensive reviews.^{2,4} Conversely, peri-implant health (PIH) refers to the absence of clinical inflammation of soft tissues, no increase in probing depth, and no further bone loss after physiological remodelling occurring during the first year of function.

Currently, periodontal probing and radiographic analysis are the primary clinical and instrumental approaches for assessing the health of dental implants.⁶ Periodontal probing determines pocket depth and eventual bleeding or suppuration, whereas radiographs can identify peri-implant bone levels. However, these traditional techniques have drawbacks that make proper assessment difficult. The absence of a peri-

^aUniversity of Trieste, Department of Chemical and Pharmaceutical Sciences, via L. Giorgeri 1, 34127 Trieste, Italy. E-mail: sfornasaro@units.it

^bMaxillofacial and Dental Surgical Clinic, Department of Medical, Surgical and Health Sciences, University of Trieste, Piazza dell'Ospitale 1, 34125 Trieste, Italy

^cRaman Spectroscopy Lab, Department of Engineering and Architecture, University of Trieste, via A. Valerio 6a, 34127 Trieste, Italy

†Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3an01438j>

odontal ligament, as well as the complexity of prosthetic designs, might make evaluation difficult. Furthermore, peri-implant mucosal seal has a lower resilience than natural teeth, resulting in easier mechanically caused bleeding on probing even in healthy implants. Although radiograph standardization and comparison with baseline references are advised, a practical implementation may be problematic. Moreover, a baseline radiograph may be not always available. Furthermore, the lack of a prediction model for peri-implantitis progression continues to be an important issue in determining prognosis.⁷ Today, in the clinical practice, it may be problematic to determine PIM or PI presence and/or progression on the base of a single examination of the affected implant. Early diagnosis and progression rate remain a great challenge, and reliable analytical methodologies are required to provide an objective assessment of implant health.

A promising approach to enhance the assessment of implant health and early detection of complications is the analysis of peri-implant crevicular fluid (PICF), sometimes described as peri-implant sulcular fluid (PISF), collected from the peri-implant sulcus around dental implants.⁸ PICF is a physiological fluid and an inflammatory exudate similar to the gingival crevicular fluid (GCF). The physiological function of PICF is to protect the peri-implant tissue and to mechanically cleanse the peri-implant sulcus or pocket. PICF's composition is similar to GCF, containing substances from the host and from microorganisms in the sub-mucosal and supra-mucosal plaque.^{9,10} Its production can increase in some physiological conditions, such as biting hard food, brushing teeth, or in case of both mechanical or chemical injury to the peri-implant soft tissues. The inflammation status increases the flow and changes the composition of the fluid,^{10–12} with variations in a wide range of biomolecules including cytokines, enzymes, inflammatory mediators, and immune cells. The analysis of PICF is a promising tool for a more reliable and sensitive assessment of implant health, facilitating timely interventions and personalized treatment strategies. For these reasons, an extensive search has been made to identify PICF components that could serve as diagnostic or prognostic markers providing critical insights into the local inflammatory and biological processes occurring around implants.^{6,11} For instance, Xu *et al.* reported that PICF in case of peri-implantitis contains higher levels of active collagenolytic matrix metalloproteinase (MMP) than GCF from similar sites of natural teeth affected by deep chronic periodontitis.¹² More recently, a 10-year retrospective investigation comparing crevicular fluid biomarkers from implants and teeth concluded that increased levels of MMP-8 and IL-1 β in PICF or GCF may be associated with inflammation around implants and teeth while lower levels of MMP-1/TIMP-1 may be an indicator of disease progression around implants.¹³ However, these proteins could not be used directly for diagnostic purposes due to low specificity and sensitivity.^{6,14} Multi-biomarker approach or untargeted metabolomics and interventional, longitudinal trials aimed to identify a unique set of biomarkers in PICF are still missing.

Due to its excellent molecular specificity, and label-free operation, surface-enhanced Raman scattering (SERS) spec-

troscopy has been identified as a promising technology for clinical monitoring.^{15,16} Utilizing SERS spectroscopy, we previously demonstrated the efficacy of analysing the molecular composition of GFC and identifying specific biochemical profiles associated with periimplantitis and gingivitis.¹⁷ In another work by d'Apuzzo *et al.*, SERS spectra of GCF were used to monitor different phases of the orthodontic treatment.¹⁸

The present study investigates the possibility to associate the molecular fingerprint of peri-implant crevicular fluid (PICF), obtained by SERS spectroscopy, to the state of peri-implant health, peri-implant mucositis, and peri-implantitis. The SERS spectra of the three conditions are compared using an in-depth statistical data analysis workflow, involving canonical powered partial least squares coupled to linear discriminant analysis (CPPLS-LDA) classification modeling, and repeated double cross-validation (r-dCV) of the classification model. This chemometrics workflow allowed us to integrate the analysis of PICF with existing clinical evaluation methods representing a potential opportunity to advance the management and long-term outcomes of dental implant patients and facilitate timely interventions and personalized treatment strategies to halt the progression of complications.

Experimental

Materials

All analytical-grade chemicals were purchased from Merck (Darmstadt, Germany) and used as received. Throughout the experiments, ultrapure deionized (DI) water with a resistivity of 18.2 Ω cm at 25 °C was obtained using a Millipore Milli-Q system (Merck, Germany). Methylcellulose PICF collection strips (Periopaper®) were purchased from Oraflow Inc. (NY, USA).

Study design and data collection

The “surface-enhanced Raman spectroscopy of crevicular fluid in the screening of periodontal and peri-implant pathologies” study was a prospective, single-center, cross-sectional diagnostic phase-II study conducted at the Maxillofacial Surgery and Stomatology Clinic of University Hospital Giuliano-Isontino of Trieste, Italy, between February 2021 and September 2022. The study was designed in accordance with the Helsinki Declaration (1964) and its subsequent amendments, as well as good clinical practice (GCP) guidelines, and it was approved by the relevant ethical committee (Comitato Etico Regionale Unico Friuli Venezia Giulia) under the identification number 2020-OS-283. Prior to signing an informed consent form, all qualified patients were provided with comprehensive information regarding the study protocol, comprising eventual surgical and prosthetic procedures, follow-up appointments, potential risks, and alternative therapeutic options. The patients provided signed informed consent for the utilization of their data for research purposes. Inclusion criteria specified that eligible patients were required to be 18 years of age or

older and meet the following conditions: (i) presence of at least one implant-supported rehabilitation, with follow-up >1 year after prosthetic loading; (ii) ASA score ≤ 2 , indicating good health status. Patients who met any of the following criteria were excluded from the study: (i) head and/or neck radiotherapy within the previous 10 years; (ii) chemotherapy; (iii) medications that may impact the immune system; (iv) immune pathologies that may affect the immune response; (v) oncologic patients; (vi) pregnancy or breast-feeding at any point during the study. A numerical code was assigned to each patient to ensure confidentiality. Both remote and near-pathological anamnesis were conducted, with a focus on variables such as age, gender, medical conditions (e.g. diabetes mellitus and osteoporosis), pharmacological treatment (e.g. use of anti-resorptive drugs and/or hormonal therapy), and smoking status (i.e. non-smokers, smokers with less than 10 cigarettes per day, and smokers with more than 10 cigarettes per day). Three calibrated operators performed all the clinical and radiographic evaluations of peri-implant health and determined the following parameters on six sites per implant (mesio-buccal, buccal, disto-buccal, mesio-lingual, lingual, and disto-lingual): plaque index (IP), bleeding on probing (BoP), and peri-implant probing depth (PPD).⁴ IP is used for the evaluation of plaque accumulation; BoP to evaluate the presence of bleeding, and PPD to record the depth of the pocket. The average of the measurements from the six sites was considered to represent the reading of the individual implant for each one of the clinical scores. Following the consensus report of workgroup 4 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions, implants that did not show clinical signs of inflammation, either bleeding and/or suppuration on gentle probing were diagnosed as healthy; implants diagnosed with PIM showed clinical signs of inflammation of the peri-implant mucosa, such as oedema and erythema, along with bleeding and/or suppuration on gentle probing, with or without an increase of probing depth compared to previous measurements, in the absence of peri-implant bone loss beyond the initial bone remodelling.^{4,19} Implants from which PI samples were collected were characterized by the presence of the same clinical signs of inflammation, together with an increase in probing depth, compared with previous recordings, and a progressive bone loss, evaluated both clinically and radiographically. In cases in which baseline data, such as probing depth and radiographic evaluation one year after prosthetic loading were not available, implants presenting a combination of clinical signs (inflammation, probing depth ≥ 6 mm, bleeding and/or suppuration on gentle probing, and radiographic bone loss >3 mm, calculated as a distance between the implant platform and the actual peri-implant bone level⁴), were considered affected by PI.

PICF sampling

The PICF samples were collected using a sterile Periopaper® strip, in accordance with the protocol reported by Papagerakis *et al.*²⁰ Prior to the collection, manual removal of any visible

plaque was performed, and the vestibular portion of the crown and the vestibular peri-implant mucosa were dried using the aspirator and the air-water syringe. A dry cotton roll was positioned in the vestibular fornix and a Periopaper® strip was inserted for 2 mm in the peri-implant sulcus for 30 seconds. Strips that exhibited visible contamination by blood, pus, or saliva were discarded. Following PICF collection, each strip was placed into its corresponding labeled 1.5 mL microfuge tube. Each tube was promptly stored at -4 °C and subsequently relocated to a -80 °C freezer until the time of analysis.

Sample preparation for SERS measurements

On the day of analysis, each strip was kept at room temperature, processed by adding a colloidal dispersion of citrated-reduced silver nanoparticles (Ag-NPs), centrifuged (10 min, 4289g), and allowed to dry at room temperature for about 15 minutes, as described previously.^{17,21} To facilitate handling, all dried strips were placed on a standard microscope slide (25 × 75 mm) that was fitted onto the portable microscope stage before measurement. SERS measurements were performed on 15 different days and over 3 different batches of Ag-NPs. A stratification of samples was conducted across different batches and days to ensure that an approximately equal number of samples from each batch and the three classes were measured on a daily basis. Thus, any dissimilarities among the classes cannot be attributed to either the day of measurement or the colloidal batch employed.

SERS instrumentation

SERS spectra of PICF samples were recorded at room temperature (22 ± 0.5 °C) utilizing a portable i-Raman Plus integrated system (BWS465785S, B&W Tek, Newark, DE, USA). The instrument was equipped with a CleanLaze® 785 nm laser that was tuned to deliver ~ 38 mW at the sample. The laser was focused on a spot with a diameter of roughly 105 μm through the BAC151B Raman video micro-sampling system, which was equipped with a 20× Olympus objective (working distance 8.8 mm, N.A. 0.25). The instrument settings were adjusted to enhance the signal and prevent any deterioration of the sample caused by the laser excitation. The spectral acquisition was performed using the BWSpec™ version 4.03_23_c (B&W Tek., Newark, DE, USA) software in the Raman shift range of 62–3202 cm^{-1} , with an average spectral resolution of 3.22 cm^{-1} , using a 10 s CCD exposure for a single accumulation. The BWSpec™ software allowed the collection of a background signal (dark) before data acquisition and its subtraction from the collected data. To compensate for intra-substrate variability, a total of 3 spectra were collected from distinct locations of the same substrate and subsequently averaged into one spectrum per sample. Wavenumber calibration was checked before and during every data collection session by recording a spectrum of 4-acetaminophenol, as a standard reference.

Data analysis

All data processing and visualization were performed within the R software environment for statistical computing and graphics (R version 4.1.2 “Bird Hippie”), building on the packages hyperSpec,²² baseline,²³ pls,²⁴ and RVAideMemoire,²⁵ on a commercially available workstation (x86_64-apple-darwin17.0 64-bit). In-house developed R scripts were used for visualization and further processing. Prior to the statistical analysis, all raw SERS spectra were preprocessed as follows: (i) the cosmic ray noise was removed, using a simple moving average algorithm proposed by Whitaker and Hayes;²⁶ (ii) spectral outliers were identified within each class of samples with the functional boxplot approach proposed by Sun and Genton²⁷ and removed from the dataset; (iii) all spectra underwent a smoothing interpolation onto an evenly spaced wave-number axis from 350 to 2000 cm^{-1} with data point spacing of 3 cm^{-1} , to simultaneously reduce noise and decrease the number of variables for samples; (iv) a baseline obtained with an asymmetric least squares (AsLS) algorithm ($\lambda = 4$ and $p = 0.005$) was fit automatically to the whole spectral range and was subtracted from each spectrum of dataset (v) total intensity vector normalization was applied to effectively normalize SERS spectra in the entire spectral region. After preprocessing, a multivariate statistical analysis, based on canonical-powered partial least squares (CPPLS) and linear discriminant analysis (LDA), was used (Fig. 1). CPPLS integrates canonical correlation analysis and the parameterization of loading weights optimized over a given interval.²⁸ In this study, cppls.fit function from the pls package²⁴ was implemented as a feature reduction step before the use of the LDA classifier. The extracted CPPLS scores were then combined with clinical scores (IP, BoP, PPD) and fed into the LDA classifier (function lda from the MASS package²⁹), after being block scaled in function of total variance of each block to balance the contributions of the two different sources during the model training. The scaling coefficients were stored in case the model needs to be applied to a new test data set. The statistical significance of the classification was evaluated by computing the classification error rate through a repeated double cross-validation

(rdCV) scheme, as suggested in ref. 30 and 31, and first implemented for SERS analysis of biofluids in ref. 32. Briefly, all samples were randomly divided into three stratified subsets (*i.e.*, respecting the proportional representation of the three implant classes in the complete data) and arranged in two nested loops of cross-validation. The first subset of samples was set aside to mimic an external test group (outer loop, 3-folds). Meanwhile, the second and third subsets were used in the inner loop (7-folds) to evaluate CPPLS dimensionality, *i.e.* the determination of the optimal number of latent variables (LVs) to be retained in the model by applying the so-called “one-standard-error” rule.³¹ To mitigate the potential influence of the composition of the folds on the final model performances, the entire procedure was repeated 200 times, with each repetition entailing a different distribution of samples within the cross-validation folds. Confusion matrices were obtained for each of the 200 repetitions by averaging the partial confusion matrices over each fold. All figures of merit (FoMs) were generated from the predictions and presented along with the upper and lower margins of the 95% credible intervals (CIs), calculated using the ci function in the bayestestR package.³³ Permutation tests with 999 randomizations were also performed using the function MVA.test from the package RVAideMemoire,²⁵ to check the reliability and the significance of the classification results through the establishment of the null distribution of the primary classification metrics and the computation of the corresponding *p*-values, using the Benjamini and Hochberg correction for multiple testing.³⁴

Results and discussion

Patient's characteristics

The clinical study included 305 implants from 118 patients, with an age at the diagnosis ranging from 31 to 87 years (median 72, IQR 50-84). Among these patients, 168 implants were from women and 136 were from men. Clinical characteristics of the patients are shown in Table 1. Groups of implants

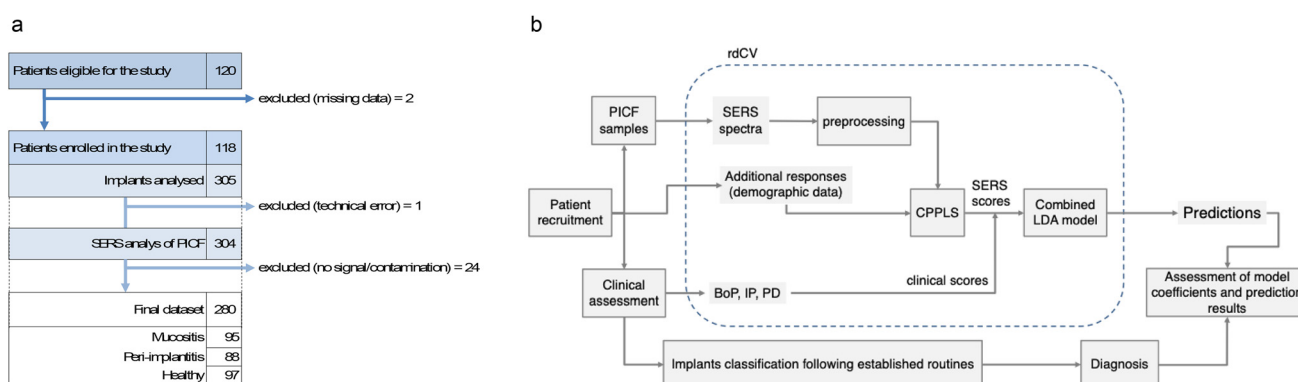


Fig. 1 Patient recruitment and data processing. a. Flow chart for patient recruitment during the study. b. Data processing and evaluation from patient recruitment to evaluation. PICF, peri-implant crevicular fluid; rdCV, repeated double cross-validation; BoP, bleeding on probing; IP, plaque index, PPD, peri-implant probing depth.

were fairly evenly distributed by age and sex, with a pronounced skew toward samples belonging to women in the group of implants diagnosed with PI. Out of the 304 implants that were analysed using SERS spectroscopy (one implant was not considered because of a technical problem during storage), 24 were excluded the study after careful data inspection. The exclusion was based on the absence of any spectral feature other than the ones attributed to the bare substrate (methylcellulose strip covered with Ag-NPs) or to the presence of contamination from other biological fluids, such as blood and purulent exudate. Both outcomes indicated the inadequacy of PICF sampling for the specific samples that were rejected. After this step, 280 samples (840 spectra), corresponding to 92% of the expected data set, were further preprocessed and analysed.

Spectral analysis

Profiles of the SERS spectra aggregated by the diagnosis class, along with their median difference spectra and the corresponding interquartile range (IQR) distributions, overlying as shaded fills, are depicted in Fig. 2a. The overall median profiles are consistent between the classes, with the prominent SERS bands located at around 494, 651, 723, 807, 927, 1005, 1128, 1215, 1326, 1443, 1581, and 1640 cm^{-1} clearly observed for each class, reflecting the overall PICF composition as detected by SERS. The PICF SERS spectra in Fig. 2a presented many similarities with those recently reported from GCF,¹⁷ displaying typical spectral features of biological fluids, such as saliva,³⁵ serum,^{36,37} and tears,³⁸ which are primarily dominated by the bands of a small number of low-molecular-weight (*i.e.* <3 kDa) metabolites. These characteristic bands can be attributed to the vibrational modes of ergothioneine (bands centered at 494, 1215, and 1443 cm^{-1}), the reduced form of glutathione (bands at 660, and 807 cm^{-1}), hypoxanthine (the band at 723 cm^{-1}), and uric acid (bands centered at 651 and 1125 cm^{-1}). To confirm this interpretation, the whole dataset of SERS spectra of PICF was fit to a linear combination of the SERS spectra of aqueous solutions of ergothioneine, glutathione hypoxanthine, and uric acid. The signal from the paper strip covered with Ag-NPs was also accounted for as the back-

ground signal. Quick and simple to implement, this technique yielded satisfactory fitting results for a qualitative assessment. The fitting procedure captured nearly all major vibrational features and their relative intensities (Fig. 2b), with a median lack-of-fit of 14% (95% CI 13–15%). The rapid immobilization of Ag-NPs on the sampling paper surface induced by the centrifugation was accompanied with an almost complete reduction in interference from large proteins, as previously reported.^{17,21} Furthermore, the presence of residual spectral contributions from vibrational bands associated with aminoacidic residues (927 cm^{-1} , $\nu(\text{C}-\text{C})$, and 1005 cm^{-1} , Phe ring breathing) is not surprising given the high abundance of small proteins, especially pro-inflammatory cytokines, within the PICF, contributing to various physiological and pathological processes.³⁹

Despite the similarities, subtle differences reflecting selected small alterations in the overall biochemical composition of the PICF associated with the two considered pathological conditions can also be observed (Fig. S1†). However, these small differences in spectral band positions and intensities were dominated by intraclass variability, making them difficult to discern by eye. Therefore, advanced chemometric data analysis and prediction algorithms were used to decipher underlying biomarker patterns for the proper interpretation and classification of spectra.

Multivariate analysis

CPPLS and LDA were used to test the ability of PICF SERS spectra to differentiate implants in PIM or PI from healthy implants. CPPLS was first used to extract diagnostically significant SERS spectral features for class discrimination. The CPPLS algorithm is designed to deliver latent variables which are optimal for discriminating the different cases, coded in a dummy response matrix describing class membership, with each class represented by a separate column and coded as either 1 or 0 to indicate belonging or not belonging to the class, respectively. A complete description of the algorithm can be found in ref. 28. When compared to standard methods for data reduction such as PCA, CPPLS has showed the capacity to extract more information from spectral data in the first components. Furthermore, this method provides the distinct advantage of integrating data from supplementary variables (*e.g.*, design factors, some kind of summary of the data or other discrete or continuous values that contain information about the data set, not to be considered predictors or responses), to stabilize predictions from difficult data and simplify the corresponding models, thus facilitating the understanding of results. Here, information about patients' age and gender was incorporated, as proposed by Ramoji *et al.*,⁴⁰ to address model bias related to patient demographics. The primary concern associated with utilizing CPPLS on spectral data pertains to the possibility of overfitting, which can lead to the development of excessively optimistic models. To address this issue, a rdCV scheme was implemented. The rdCV employs an iterative methodology in which all samples are sequentially assigned to either the calibration or validation set.

Table 1 Clinical characterization of patients

Characteristics	PIM	PI	PIH
Number of patients, <i>n</i> (%)	105 (35)	95 (31)	104 (34)
Sex, male, <i>n</i> (%)	48 (35)	23 (17)	65 (48)
Age, median (IQR)	74 (65–79)	73 (64–79)	69 (61–75)
Non-smokers, <i>n</i> (%)	86 (36)	74 (30)	82 (34)
Light smokers ^a , <i>n</i> (%)	13 (31)	13 (31)	16 (38)
Heavy smokers ^b , <i>n</i> (%)	6 (30)	8 (40)	6 (30)
Anti-resorptives, <i>n</i> (%)	15 (56)	3 (11)	9 (33)
Hormonal therapy, <i>n</i> (%)	29 (44)	25 (38)	12 (18)
Not taking drugs, <i>n</i> (%)	61 (29)	67 (31)	83 (40)
Diabetes mellitus, <i>n</i> (%)	21 (40)	14 (26)	18 (34)
Osteoporosis, <i>n</i> (%)	26 (49)	22 (42)	5 (9)
No pathologies, <i>n</i> (%)	58 (29)	59 (30)	81 (41)

IQR, interquartile range. ^a <10 cigarettes per day. ^b >20 cigarettes day.

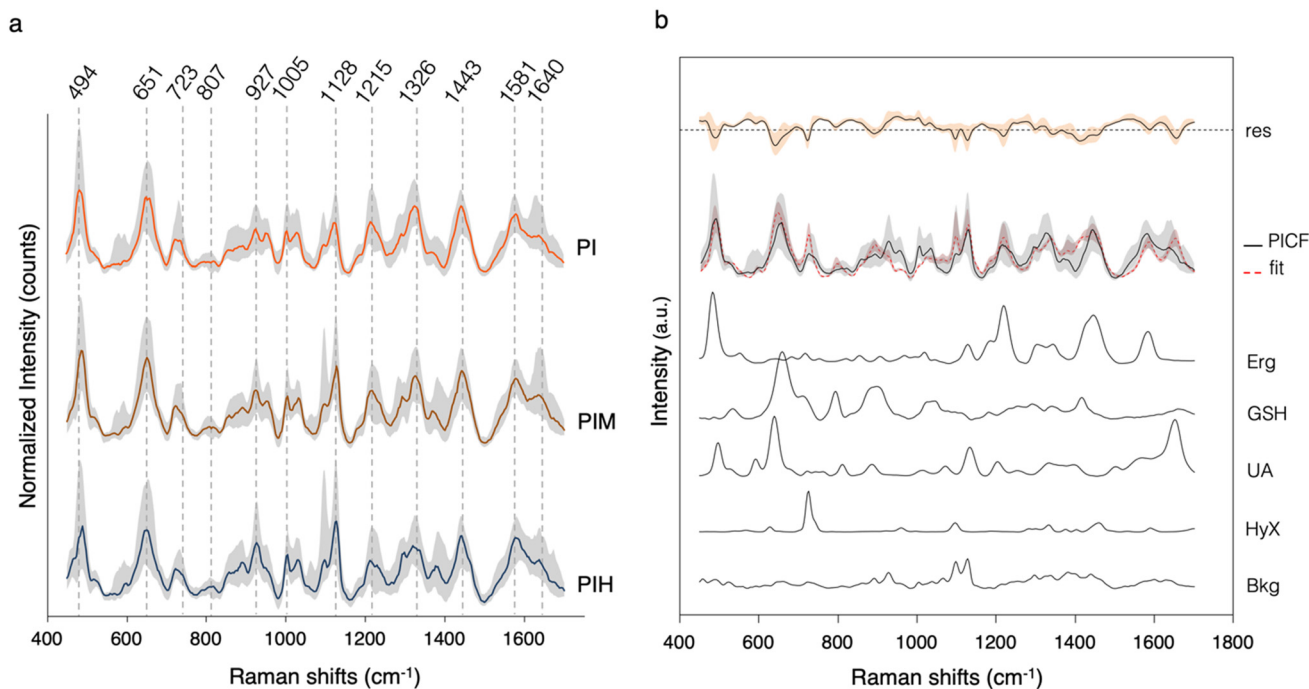


Fig. 2 SERS spectra of PICF. a. Median (solid lines) and interquartile (IQR) range (shaded areas) of the baseline-corrected and intensity-normalized SERS spectra for PI, PIM, and PIH samples. The spectra are shifted on the y-axis to avoid overlap. The key SERS bands were marked by dashed lines. b. Comparison of SERS spectra of PICF and some model compounds. The dashed red line represents the best-fit spectrum (fit), *i.e.*, the spectral sum of all components as displayed in this figure: ERG, ergothioneine; GSH, reduced glutathione; UA, uric acid; HyX, hypoxanthine; Bkg, a naked Periopaper strip covered with Ag-NPs. The best-fit spectrum was subtracted from each PICF spectrum to obtain the residuals' distribution (res). Lack-of fit = 14% (95%CI 13–15%); $r^2 = 0.99$ (95% CI 0.98–0.99).

The rdCV methodology is designed so that every optimization and validation iteration is executed on a distinct and independent dataset. Since no data from the test set is incorporated into the model's training phase, overfitting is prevented. The optimal number of latent variables (LVs) was chosen in the inner rdCV loop (Fig. S2b†) and combined with clinical scores (PI, BoP, PPD) to produce a column-augmented matrix to be used as input for the LDA. In the outer rdCV loop, the performance of each optimized model generated by the rdCV was validated by comparing the predictions to the actual classes of an independent test set. Fig. S2a† displays the cross-validation error curves for all models obtained in the inner loop, demonstrating that the classification error rate does not improve when using a more complex model (up to 10 LVs), which explains 48% of the variance (compared to 46% in the one-component model). These data imply that the LVs after the first are ineffective at distinguishing between the three classes. Similarly, only variable loading weights along the first LV can be considered significant (Fig. S2c†), with nearly all the 95% confidence intervals for the pseudo bands (shaded regions) starting to cross zero from the second LV onward.⁴¹ This indicates that only the spectral bands identifiable along the first LV maintained consistent importance throughout the entirety of the rdCV process, irrespective of the samples selected at each step of validation.

The corresponding CPPLS coefficients' vectors, shown in Fig. 3, reflected the SERS spectral features responsible for class

differentiation in terms of band intensities and band positions, *i.e.*, changes in relative concentration or composition. A positive coefficient indicates a positive relationship with the response (*i.e.*: class membership), and conversely for a negative coefficient. It is evident that the coefficients for the PIM and PI classes were nearly identical, differing only in terms of general intensity, and both appear to be opposite to those of the healthy implant class. Different regions within the coefficients vector can be roughly assigned to the same bands associated with metabolites involved in oxidative and inflammatory tissue responses. The positive peaks in the coefficients for PIM and PI can be interpreted as glutathione, ergothioneine, and hypoxanthine bands. Ergothioneine and glutathione are known to accumulate in response to oxidative stress, acting as antioxidants and protectors of tissue integrity.^{42,43} Similarly, the increased presence of hypoxanthine, a purine metabolite, has been linked to ongoing inflammatory processes in periodontal disease.⁴⁴ On the other hand, the positive peaks in the PIH class coefficients vector seem to be correlated with the background signal of the paper, probably indicative of a lower presence of PICF in healthy implants.⁴⁵ The amount of PICF produced at a site significantly increases with the severity of gingival inflammation, because it is directly related to the augmentation of the vascular permeability and ulceration of the pocket epithelium.⁴⁶ A plausible conveyance of the two previous statements is that the

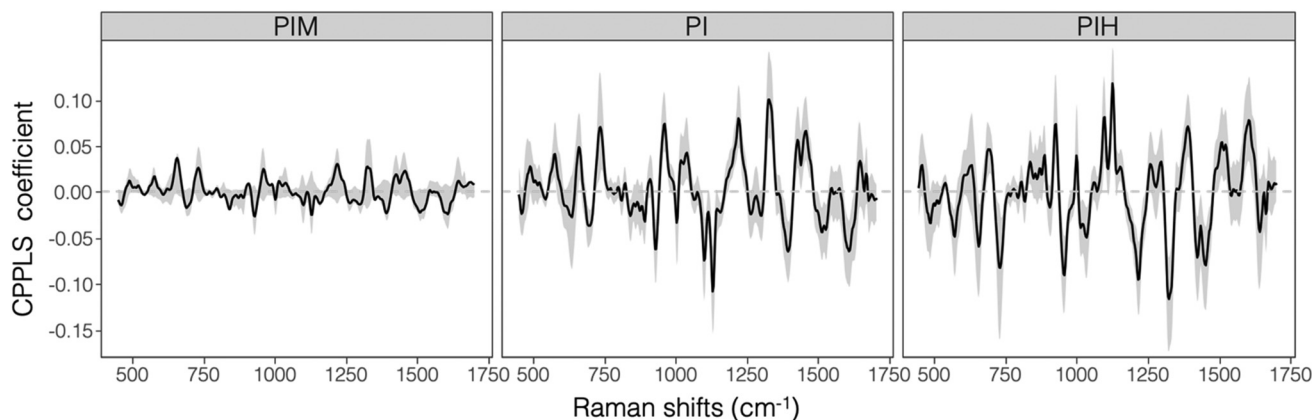


Fig. 3 Canonical powered partial least squares (CPPLS) coefficients for the repeated double cross-validation (rdCV) model. PIM, peri-implant mucositis; PI, peri-implantitis; PHI, peri-implant health.

distinct clinical presentations and outcomes observed in these two conditions (PI and PIM) may be just the results of variations in the severity or in the temporal progression of a shared biochemical phenomenon, influenced by factors such as the host's immune response and the local microenvironment. Interestingly, the histopathologic and clinical conditions leading to the conversion from one condition to the other are not completely understood.⁵ The coefficients' 95% CI (estimated non-parametrically from the distributions obtained by rdCV) are also graphically represented in Fig. 3. If the shaded regions include zero, it indicates that the coefficient is not statistically significant, and the relationship between the spectral features and the sample class may be due to random chance rather than a true association. The confidence intervals showed that the iterative rdCV process, while producing changes in their absolute values, does not systematically alter their sign, demonstrating that the variables maintain their positive or negative correlation with each class. The coefficients' vector for the PIM samples exhibited lower significance compared to the vectors for the other two classes. However, it is important to interpret this lower significance of the coefficients in the PIM samples cautiously. While the statistical significance may be lower, it does not necessarily imply that the spectral differences between PIM and the other classes are insignificant in a clinical sense. If we consider PIM as an early stage of PI with milder inflammation and tissue changes, it makes sense that the biochemical alterations in PIM were less pronounced compared to the more advanced and severe PI cases, leading to greater variability in the spectral fingerprints, thus reducing the statistical significance of the associations. The LDA was successfully employed as a robust classifier by projecting the scores into a discriminant space along the canonical variable and afterward for a prediction of class membership probabilities. To facilitate the interpretation and visualization of the CPPLS-LDA models, we adopted the canonical variates approach, utilizing predicted values obtained from rdCV. This approach provides a natural and comprehensive way to understand the discriminatory power of the model.

Fig. 4a shows projections along the first canonical variates, showcasing the separation between the different sample classes. It is evident that the healthy, PHI, implants can be distinctly separated from the ones with PIM and PI along this direction. As already discussed, the spectral features captured by the CPPLS almost perfectly differentiate healthy implants from those affected with PI, with a slightly worse discriminative power for the implants with PIM. Moreover, complete discrimination between the PIM and PI classes was not always feasible, with a classification error rate for these two classes between 8 and 10%. A complete picture can be seen in the confusion matrix in Fig. 4b. Taking advantage of the r-dCV procedure, figures of merit (FoM) such as accuracy, precision, sensitivity, and specificity were calculated from the confusion matrix of each optimized model, yielding a distribution of values for each FoM from which CIs were calculated (Table S1†). Overall balanced accuracy was about 92%. Implants with PI were recognized with 90% sensitivity (89–91%) and 96% specificity (95–97%), and implants with PIM were recognized with 89% sensitivity (87–92%) and 92% specificity (91–92%). It should be stressed that, during the rdCV iterative process, almost all samples maintained their original classification, as depicted by the extremely narrow 95% CIs in the confusion matrix. To further assess the statistical significance of these results, they were compared to those obtained from a validation of permuted data (*i.e.*, permutation test) in which the class labels were randomly reassigned, and subsequently, a new classification model is computed for each iteration. If the performance of the latest models derived from permuted class labels is consistently and significantly lower than that of the original model, it can be inferred that the original model is robust and reliable. In this case (Fig. 5), 999 permutations resulted in classification error rates around 70%, extremely far from the model rate (10%). None had a classification error rate lower than 55%, leading to a p -value = 0.001 after FDR. In practice, no random sample assignment is even vaguely able to simulate a correct classification, as expected for unbiased models.

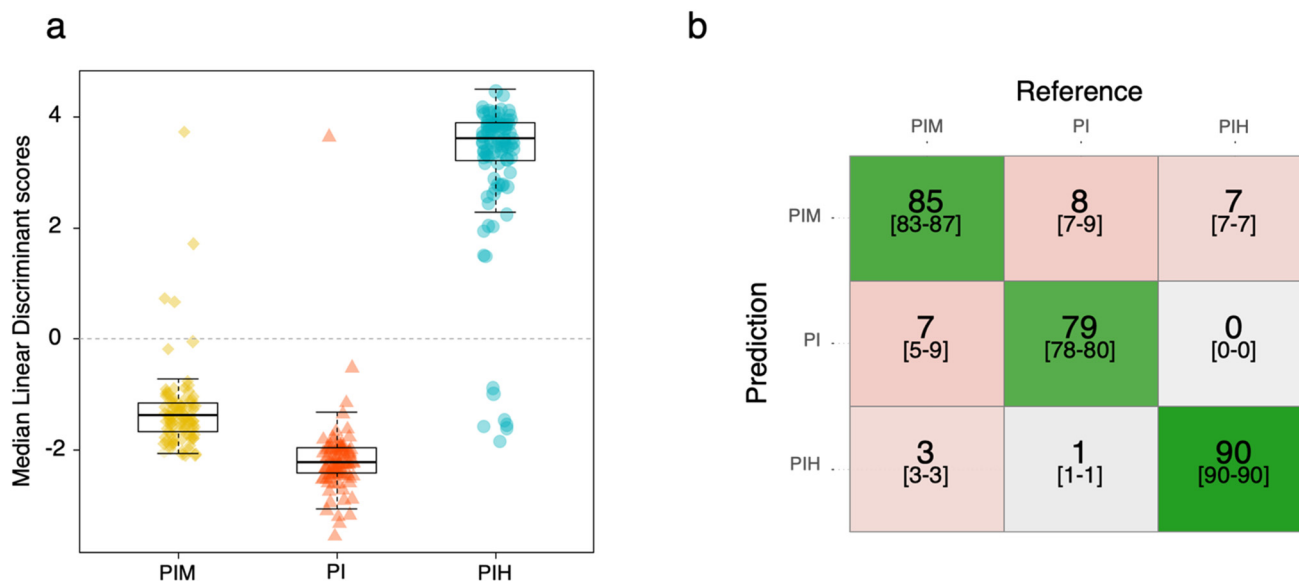


Fig. 4 a. Medians of the LD scores for each sample, calculated over the optimized models from the rdCV. b. Confusion matrix resulting from the predictions of the optimized rdCV model showing median and 95% CI after 200 repetitions.

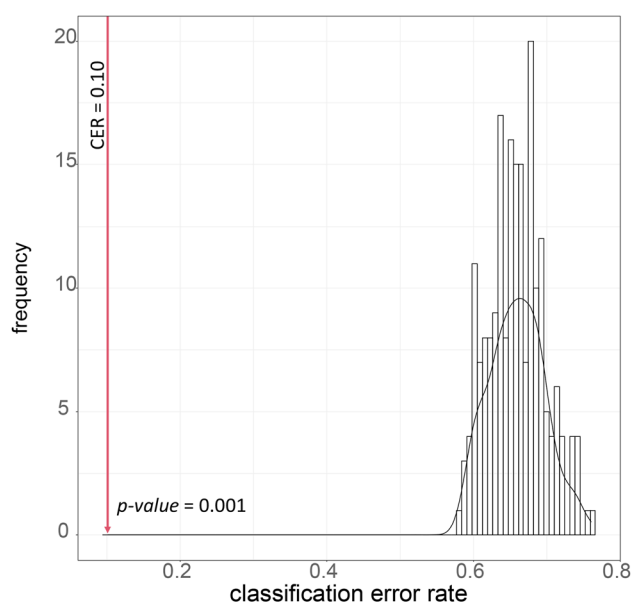


Fig. 5 Permutation test. Class prediction result based on rdCV model compared to the permuted ($n = 999$) data. Classification error rate for the rdCV model is reported with a red arrow.

Although this study showed promising results and represents the first study with a medium sample size to explore the diagnostic ability of SERS analysis of PICF in a clinical cohort, several limitations are present. A cross-sectional study does not allow any claim of causality, thus only future prospective studies with a larger multicenter dataset would provide conclusive insights into the utility of SERS analysis of PICF in diagnosing and monitoring peri-implant diseases.

Conclusions

Surface-enhanced Raman scattering (SERS) of peri-implant crevicular fluid offers a promising approach for visualizing the underlying inflammatory process linked with peri-implant mucositis and peri-implantitis. This methodology presents the advantage of practically null invasivity for the patient, minimal sample preparation, enabling rapid assessment of the implant's status with high accuracy and relatively negligible costs. The outcomes of this study highlight the potential utility of SERS spectroscopy coupled with chemometrics as a diagnostic and prognostic tool in oral implantology, capable of differentiating between overlapping conditions. This can have significant implications for enhancing the quality of patient care. Looking towards the future, one particularly notable approach would involve the integration of the quantification of pro-inflammatory cytokines such as IL-17A, IL-6, IL-15, and matrix metalloproteinase 8 (MMP8), associated with the development of peri-implantitis,⁴⁷ complementing the SERS biochemical fingerprint of PICF, and affording a more comprehensive understanding of the inflammatory environment. The resulting complex set of heterogeneous and complex data will benefit from the use of advanced machine learning techniques, such as convolutional neural networks and deep learning. This comprehensive strategy may allow for improved association of SERS-derived molecular information with specific inflammatory status, hence improving diagnostic accuracy. Furthermore, a future clinical study could incorporate a longitudinal component, involving the collection of samples at regular intervals to closely monitor and evaluate the response to therapeutic interventions in cases of mucositis, a condition inherently characterized by its reversible nature. By tracking the progress or evolution of the condition towards peri-implan-

titis, this study design would provide critical insights into the efficacy of the therapeutic regimen and guide further refinement of treatment strategies.

Author contributions

Conceptualization, S. F. and F. B.; methodology, S. F., A. B., and F. B.; formal analysis, S. F.; investigation, S. F., A. R., F. F., M. I., and G. P.; resources, A. B., V. S., A. R., F. B., C. S.; data curation, A. R. G. P. and S. F.; writing – original draft preparation, S. F., F. B.; writing – review and editing, S. F., A. R., F. F., M. I., G. P., C. S., V. S., A. B., R. D. L., F. B.; visualization, S. F. and A. B.; supervision, F. B., A. R.; project administration, F. B.; funding acquisition, F. B., R. D. L. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

References

- 1 S. Windael, B. Collaert, S. De Buyser, H. De Bruyn and S. Vervaeke, *Clin. Implant Dent. Relat. Res.*, 2021, **23**, 298–308.
- 2 J. Derks and C. Tomasi, *J. Clin. Periodontol.*, 2015, **42**, S158–S171.
- 3 L. J. A. Heitz-Mayfield and G. E. Salvi, *J. Periodontol.*, 2018, **89**, S257–S266.
- 4 T. Berglundh, G. Armitage, M. G. Araujo, G. Avila-Ortiz, J. Blanco, P. M. Camargo, S. Chen, D. Cochran, J. Derks, E. Figuero, C. H. F. Hammerle, L. J. A. Heitz-Mayfield, G. Huynh-Ba, V. Iacono, K.-T. Koo, F. Lambert, L. McCauley, M. Quirynen, S. Renvert, G. E. Salvi, F. Schwarz, D. Tarnow, C. Tomasi, H.-L. Wang and N. Zitzmann, *J. Periodontol.*, 2018, **89**, S313–S318.
- 5 F. Schwarz, J. Derks, A. Monje and H.-L. Wang, *J. Periodontol.*, 2018, **89**, S267–S290.
- 6 H. Alassy, P. Parachuru and L. Wolff, *Diagnostics*, 2019, **9**, 214.
- 7 S. Renvert, G. R. Persson, F. Q. Pirih and P. M. Camargo, *J. Periodontol.*, 2018, **89**, S304–S312.
- 8 F. Faot, G. G. Nascimento, A. M. Bielemann, T. D. Campão, F. R. M. Leite and M. Quirynen, *J. Periodontol.*, 2015, **86**, 631–645.
- 9 R. AlRowis, H. S. AlMoharib, A. AlMubarak, J. Bhaskardoss, R. S. Preethanath and S. Anil, *J. Int. Oral Health*, 2014, **6**(5), 126–135.
- 10 P. M. Duarte, C. R. Serrão, T. S. Miranda, L. C. S. Zanatta, M. F. Bastos, M. Faveri, L. C. Figueiredo and M. Feres, *J. Periodontal Res.*, 2016, **51**, 689–698.
- 11 E. G. Kaklamanos and L. Tsalikis, *J. Int. Acad. Periodontol.*, 2002, **4**, 49–59.
- 12 L. Xu, Z. Yu, H.-M. Lee, M. S. Wolff, L. M. Golub, T. Sorsa and H. Kuula, *Acta Odontol. Scand.*, 2008, **66**, 219–224.
- 13 C. A. Ramseier, S. Eick, C. Brönnimann, D. Buser, U. Brägger and G. E. Salvi, *Clin. Oral Implants Res.*, 2016, **27**, 211–217.
- 14 M. Sanz, B. Noguerol, I. Sanz-Sanchez, C. H. F. Hammerle, H. Schliephake, F. Renouard, A. Sicilia, Steering Committee. European Association for Osseointegration, L. Cordaro, R. Jung, B. Klinge, P. Valentini, G. Alcoforado, T. Ornekol, B. Pjetursson, I. Sailer, I. Rochietta, J. Manuel Navarro, L. Heitz-Mayfield and H. Francisco, *Clin. Oral Implants Res.*, 2019, **30**, 476–486.
- 15 M. Chisanga, H. Muhamadali, D. Ellis and R. Goodacre, *Appl. Sci.*, 2019, **9**, 1163.
- 16 J. Plou, P. S. Valera, I. García, C. D. L. de Albuquerque, A. Carracedo and L. M. Liz-Marzán, *ACS Photonics*, 2022, **9**(2), 333–350.
- 17 S. Fornasaro, F. Berton, C. Stacchi, F. Farina, A. Esposito, V. Sergo, R. Di Lenarda and A. Bonifacio, *Analyst*, 2021, **146**, 1464–1471.
- 18 F. d'Apuzzo, L. Perillo, I. Delfino and M. Portaccio, *J. Biomed. Opt.*, 2017, **22**, 1.
- 19 J. G. Caton, G. Armitage, T. Berglundh, I. L. C. Chapple, S. Jepsen, K. S. Kornman, B. L. Mealey, P. N. Papapanou, M. Sanz and M. S. Tonetti, *J. Clin. Periodontol.*, 2018, **45**, S1–S8.
- 20 P. Papagerakis, L. Zheng, D. Kim, R. Said, A. A. Ehlert, K. K. M. Chung and S. Papagerakis, in *Odontogenesis. Methods in molecular biology*, Humana Press, New York, NY, 2019, vol. 1922, pp. 549–562.
- 21 A. Esposito, A. Bonifacio, V. Sergo and S. Fornasaro, *Biosensors*, 2021, **11**, 467.
- 22 C. Beleites and V. Sergo, *R package version 0.99-20200527*, <https://github.com/cbeleites/hyperSpec>.
- 23 K. H. Liland, T. Almøy and B.-H. Mevik, *Appl. Spectrosc.*, 2010, **64**, 1007–1016.
- 24 K. H. Liland, B.-H. Mevik and R. Wehrens, <https://CRAN.R-project.org/package=pls>.
- 25 M. Herve, <https://CRAN.R-project.org/package=RVAideMemoire>.
- 26 D. A. Whitaker and K. Hayes, *Chemom. Intell. Lab. Syst.*, 2018, **179**, 82–84.
- 27 Y. Sun and M. G. Genton, *J. Comput. Graph. Stat.*, 2011, **20**, 316–334.
- 28 U. G. Indahl, K. H. Liland and T. Naes, *J. Chemom.*, 2009, **23**, 495–504.
- 29 W. N. Venables, B. D. Ripley and W. N. Venables, *Modern applied statistics with S*, Springer, New York, 4th edn, 2002.
- 30 J. A. Westerhuis, H. C. J. Hoefsloot, S. Smit, D. J. Vis, A. K. Smilde, E. J. J. Van Velzen, J. P. M. Van Duynhoven and F. A. Van Dorsten, *Metabolomics*, 2008, **4**, 81–89.
- 31 P. Filzmoser, B. Liebmann and K. Varmuza, *J. Chemom.*, 2009, **23**, 160–171.
- 32 E. Gurian, A. Di Silvestre, E. Mitri, D. Pascut, C. Tiribelli, M. Giuffrè, L. S. Crocè, V. Sergo and A. Bonifacio, *Anal. Bioanal. Chem.*, 2020, **413**, 1303–1312.

- 33 D. Makowski, M. Ben-Shachar and D. Lüdecke, *J. Open Source Software*, 2019, **4**(40), 1541.
- 34 Y. Benjamini and Y. Hochberg, *J. R. Stat. Soc., Ser. B (Methodological)*, 1995, **57**, 289–300.
- 35 S. Chen, H. Wu, C. Chen, D. Wang, Y. Yang, Z. Zhou, R. Zhu, X. He, Y. Pan and C. Li, *Spectrochim. Acta, Part A*, 2023, **288**, 122150.
- 36 W. R. Premasiri, J. C. Lee and L. D. Ziegler, *J. Phys. Chem. B*, 2012, **116**, 9376–9386.
- 37 A. Bonifacio, S. Dalla Marta, R. Spizzo, S. Cervo, A. Steffan, A. Colombatti and V. Sergo, *Anal. Bioanal. Chem.*, 2014, **406**, 2355–2365.
- 38 P. Hu, X.-S. Zheng, C. Zong, M.-H. Li, L.-Y. Zhang, W. Li and B. Ren, *J. Raman Spectrosc.*, 2014, **45**, 565–573.
- 39 T. Halstenbach, K. Nelson, G. Iglhaut, O. Schilling and T. Fretwurst, *J. Periodontol.*, 2023, **94**(7), 835–847.
- 40 A. Ramoji, D. Thomas-Rüddel, O. Ryabchykov, M. Bauer, N. Arend, E. J. Giamarellos-Bourboulis, J. Eugen-Olsen, M. Kiehntopf, T. Bocklitz, J. Popp, F. Bloos and U. Neugebauer, *Crit. Care Explor.*, 2021, **3**, e0394.
- 41 S. Fornasaro, C. Beleites, V. Sergo and A. Bonifacio, in *SERS for Point-Of-care and Clinical Applications*, Elsevier, 2022, pp. 1–51.
- 42 I. K. Cheah and B. Halliwell, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2012, **1822**, 784–793.
- 43 S. P. Patel, A. R. Pradeep and S. Chowdhry, *Arch. Oral Biol.*, 2009, **54**, 543–548.
- 44 V. M. Barnes, R. Teles, H. M. Trivedi, W. Devizio, T. Xu, M. W. Mitchell, M. V. Milburn and L. Guo, *J. Dent. Res.*, 2009, **88**, 851–855.
- 45 L. Bevilacqua, M. De Biasi, M. G. Lorenzon, C. Frattini and D. Angerame, *Open Dent. J.*, 2016, **10**, 131–138.
- 46 S. Lachmann, E. Kimmerle-Müller, D. Axmann, L. Scheideler, H. Weber and R. Haas, *Clin. Oral Implants Res.*, 2007, **18**, 212–223.
- 47 L. Song, J. Jiang, J. Li, C. Zhou, Y. Chen, H. Lu and F. He, *J. Clin. Med.*, 2022, **11**, 5817.