Methods in Raman spectroscopy for saliva studies – a review

Mike Hardy, Liam Kelleher, Paulo de Carvalho Gomes, Emma Buchan, Hin On Martin Chu & Pola Goldberg Oppenheimer

To cite this article: Mike Hardy, Liam Kelleher, Paulo de Carvalho Gomes, Emma Buchan, Hin On Martin Chu & Pola Goldberg Oppenheimer (2022) Methods in Raman spectroscopy for saliva studies – a review, Applied Spectroscopy Reviews, 57:3, 177-233, DOI: 10.1080/05704928.2021.1969944

To link to this article: https://doi.org/10.1080/05704928.2021.1969944
Methods in Raman spectroscopy for saliva studies – a review

Mike Hardya,b, Liam Kellehera,c, Paulo de Carvalho Gomesa, Emma Buchana, Hin On Martin Chua, and Pola Goldberg Oppenheimera,d

aSchool of Chemical Engineering, College of Engineering and Physical Sciences, University of Birmingham, Birmingham, UK; bInstitute for Global Food Security, School of Biological Sciences, Queen’s University Belfast, Belfast, UK; cSchool of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK; dInstitute of Translational Medicine, Healthcare Technologies Institute, Birmingham, UK

ABSTRACT

The use of Raman spectroscopy combined with saliva is an exciting emerging spectroscopy-biofluid combination. In this review, we summarize current methods employed in such studies, in particular the collection, pretreatment, and storage of saliva, as well as measurement procedures and Raman parameters used. Given the need for sensitive detection, surface-enhanced Raman methods are also surveyed, alongside chemometric techniques. A meta-analysis of variables is compiled. We observe a wide range of approaches and conclude that standardization of methods and progress to more extensive validation Raman-saliva studies is necessary. Nevertheless, the studies show tremendous promise toward the improvement of speed, diagnostic accuracy, and portable device possibilities in applications such as healthcare, law enforcement, and forensics.

KEYWORDS

Raman spectroscopy; SERS; saliva; medical diagnostics; drugs

1. Introduction

1.1. Overview

The rapid and accurate detection of substances in biofluids is of paramount importance to healthcare, law enforcement, and forensics. At present, blood and urine are commonly acquired fluids and can be analyzed via various means (assays, high-performance liquid chromatography, mass spectrometry). However, other biofluids are being investigated, which may have different accessibility, risk, and content profiles than blood or urine. Optical methods show promise in biosensing, and specifically, there is an increasing interest in spectroscopic means that may confer benefits of speed, increased accuracy, and multiplexing. Alongside this, the development of smaller instruments needing low sample volumes for benchtop clinical or bedside use, or in-field applications at a roadside accident or a crime scene, is desirable.
One especially exciting combination of biofluid and spectroscopy is that of saliva and Raman spectroscopy. Herein, we provide a detailed review of the experimental design and methodology in Raman-saliva studies with a specific highlight on the use of surface-enhanced Raman spectroscopy (SERS). This review will highlight an important question as to whether researchers in Raman-saliva studies employ consistent means in their experimental methods, in terms of collection, processing, storage, measurement, and analysis, and whether these methods are suitable for purpose (Figure 1).

A lack of standardization in biomarker studies is evident in general. Poste[7] underlines the broad nature of the inquiry by raising questions on uniformity of practice and analysis pertinent to many different fields, suggesting that “lack of standardization, not only can affect the validity of the conclusion with an individual study, but also clearly impacts the meaningfulness of study comparisons.” Therefore, an increased emphasis on experimental approaches over application is timely, fueled by the need to address inconsistencies within a Raman-saliva context. Congruently, Surowiec et al.[8] has recently underlined the importance of experimental design in the analysis of complex mixtures to maximize useful data extraction and, subsequently, the knowledge gained from such. Details on further biofluids, alternative spectroscopic approaches, and applications of saliva as a potential diagnostic fluid are broadly reviewed in the literature.[9–19]

1.2. Why saliva?

1.2.1. Biofluids and saliva function

Biofluid analysis has become a key area of research in recent years, with an increasing awareness that many biomarkers and drugs may be detectable as indicators for disease or drug consumption.[16,19] The promise of saliva was noted in an early review by
Pfaffe et al., observing that while technologies had the required analytical sensitivity to benefit from saliva as a diagnostic medium, these had not yet been integrated into widespread clinical practice,[20] despite some appearance in simple home-testing kits.[21] More robust biomarker identification, validation, and disease association studies are needed.[20]

Saliva is an extracellular fluid produced by the salivary glands in mammals. The main bodily function of saliva is in lubrication, which is necessary for chewing, swallowing, and speech, but it also serves a purpose in pH regulation and for oral hygiene,[20,22–26] while calcium and phosphates in saliva contribute to enamel remineralization.[24]

1.2.2. Saliva composition
Saliva is mostly comprised of water and may take a frothy appearance. There are three main salivary glands in the human mouth: submandibular, sublingual, and parotid, each with its own rheological properties, alongside hundreds of minor glands.1 The major proteins in saliva are proline-rich peptides (PRPs), the glycoprotein α-amylase, and the much larger mucins, the three of which cumulatively constitute almost 80% of salivary proteins.[20] Up to 70% of salivary flow comes from the submandibular gland.[25] Saliva also contains extracellular vesicles (EVs) and lipids, and some non-salivary gland constituents such as epithelial cells, micro-organisms, and food remnants.[20] Where non-salivary gland constituents are considered, the term “oral fluid” might be more appropriate, nonetheless, “saliva” is frequently used in the literature to refer to the ensemble. Salivary constitution can change depending on hormonal and psychological effects, as well as physical exercise, oral hygiene, and whether the production of the fluid is stimulated or unstimulated. Body posture, i.e., lying down or standing up, and ambient light also affects salivary production, where darkness is reported to lower salivary flow by up to 40%.[25]

Moreover, saliva composition can vary greatly depending on collection methods and the time of day, i.e., subject to circadian rhythm, and even yearly seasons.[22,24,27–30] In the daily case, this may be simply as a result of different relative flow rates of various salivary glands.[24,28] It has been noted that while salivary flow peaks in the late afternoon in healthy subjects, compositional variations need to be considered separately, with, for instance, peaks in sodium and chloride levels occurring in the morning.[22] Also, differential flow rates are observed in certain diseases such as in Sjögren’s syndrome, an autoimmune condition characterized by epithelial cell destruction, where the flow from the submandibular and sublingual glands is lower than in healthy subjects hence affecting salivary composition.[9] Head and neck irradiation and HIV may produce similar effects.[24,31] Circadian variations are not exclusive to saliva and have been noted in serum and cerebral spinal fluid (CSF).[6] Gland-specific collection is possible, however, it is cumbersome and often done by cannulation.[20,32,33] When stimulated, saliva amount is related to the gland size, whereas unstimulated is not.[22] Gland size has been suggested to be the only reason for salivary variance noted between male and female subjects due to the mean gland size difference, although there are opposing views in the literature.[34,35] Similarly, there is a debate on the effect of subject body mass on the observed saliva composition in boys.[36,37]
In light of these salivary variations, researchers frequently conduct preliminary studies that include spectroscopic measurements combined with man-made saliva-like formulations and analyte spiking.\textsuperscript{[38–41]} Artificial saliva is a prevalent choice over actual human saliva in studies focusing on dental materials, which simulate an oral environment. Typical components of artificial saliva formulations contain mucin, which produces the prominent amide bands when measured with vibrational spectroscopies. However, despite the various existing formulations of artificial saliva, there is minimal guidance on standardization.\textsuperscript{[42]} Notably, Ionta et al.\textsuperscript{[42]} has reported the effect of different formulations of artificial saliva on enamel mineralization, demonstrating remineralization potential despite variations, including no effect from added mucin.

### 1.2.3. Saliva versus other biofluids

The use of saliva as a diagnostic biofluid presents various benefits.\textsuperscript{[43]} Saliva is easier to acquire than blood,\textsuperscript{[44,45]} thus enabling faster and less intrusive acquisition that may be performed without extensive training,\textsuperscript{[11]} pertinent in the potential use for portable point-of-care (PoC) medical devices.\textsuperscript{[46]} Further, there is a negligible risk of infection to the donor and a lower risk to the handler compared to blood.\textsuperscript{[6,11]} The ability to supervise during the collection of saliva minimizes the chance of adulteration, in contrast to urine, which may be necessary for sports-doping monitoring.\textsuperscript{[47]} At the roadside tests, immediate urine and blood collection pose significant challenges. Compared to urine, saliva often contains the parent compounds, whereas urine contains mainly metabolites.\textsuperscript{[47]} Moreover, for some substances, saliva may have a further advantage whereby the analytical sensitivity increases because the requisite substance has been orally ingested, as often evident with illicit drugs, however, analysis of oral residues, present in varying amounts, may impair quantitative analysis.\textsuperscript{[48]}

A positive correlation exists between multiple biomolecules detectable in serum and saliva due to the transfer of material \textit{via} the salivary duct across a thin layer of epithelial cells by passive diffusion, active transport, or extracellular ultra-filtration.\textsuperscript{[9,20]} Pfaffe et al.\textsuperscript{[20]} has compiled a comprehensive list of common biomolecules, which may be detected in both blood and saliva, and the clinically relevant ranges, highlighting indicators for oral and breast cancers as well as cardiovascular diseases, amongst others. The precise understanding of the movement of material from blood to saliva is still a matter of debate, as underlined by a model employed by Dadas et al., who study brain-derived proteomic biomarkers, noting the selective bias toward low molecular weight at the blood-saliva barrier.\textsuperscript{[49]} Inscore\textsuperscript{[50]} observes that illicit drugs in saliva may be in comparable concentrations to blood.

Although saliva tests in medicine are a routine practice, spectroscopic saliva studies in healthcare have not progressed much beyond the preliminary investigations into specific diseases, which seek to identify abnormal biomarkers.\textsuperscript{[9,31]} Historically, saliva studies have presented small sample numbers and have lacked validation, having not met diagnostic criteria in terms of sensitivity and specificity. Biomarker identification is attempted amongst a myriad of changing organic species, including body secretion products, putrefaction products, and lipids.\textsuperscript{[9]} Bonassi et al.\textsuperscript{[52]} has noted that biomarker identification should be preferably as near to the causal pathway of the disease as possible. Despite an array of different proteins present, the salivary composition is not as
complex as that of blood serum,\textsuperscript{[9]} which expresses as many as 10\textsuperscript{5} different proteins over a dynamic concentration range spanning 12 orders of magnitude.\textsuperscript{[53]} Moreover, total salivary protein content is relatively low, indicating that protein binding with other constituents is less likely to occur. This, for instance, means that any illicit drug compounds may exist as unbound molecules.\textsuperscript{[48]}

Furthermore, saliva has recently been highlighted as a potential biofluid for COVID-19 diagnostics,\textsuperscript{[54–57]} directly linked to SARS-COV-2 virus spread,\textsuperscript{[58]} and has been employed for detection of the novel coronavirus with at least comparable sensitivity to a nasopharyngeal swab test during the course of patients’ hospitalization.\textsuperscript{[59]}

\section*{1.3. Why Raman spectroscopy?}

\subsection*{1.3.1. Current analysis methods}

In many disease cases, biopsies and histopathological analyses can be performed, but the procedure is time-consuming, invasive, and may risk infection for patients.\textsuperscript{[60,61]} Furthermore, biopsies are often performed later in the diagnostic course and morphological or structural abnormalities may not be apparent in early pathologies.\textsuperscript{[62]} Other approaches, such as enzyme-linked immunosorbent assay (ELISA) or high-performance liquid chromatography (HPLC), are relatively slow and require skilled users. Mass spectrometry (MS) has similar issues and can suffer from non-universal ionization efficiency and ion-suppression.\textsuperscript{[60,63,64]} Moreover, MS precludes portable (handheld) analysis without significant detriment to analytical performance.\textsuperscript{[64]} Electrochemical sensors are popular, however, the reliability of anodic/cathodic peak analysis may be questionable.\textsuperscript{[65]} This has led to an interest in less invasive optical diagnostics such as optical coherence tomography (OCT).\textsuperscript{[60,62]}

\subsection*{1.3.2. Optical biosensing}

Biosensing requires detecting a panel of compounds, preferably simultaneously, rapidly, and reproducibly with high analytical selectivity and sensitivity.\textsuperscript{[66]} Most competitive biosensing solutions to date combine chemical transduction i.e., surface functionalization, and optical sensing e.g., phase change monitoring. The state-of-the-art techniques include surface plasmon resonance (SPR) and the related technique of grating coupled waveguide (GCW) interferometry,\textsuperscript{[66]} which operate based on a change in the refractive index of the aqueous medium where the analyte binds to an affinity molecule (antibodies, aptamers, small molecules, and polymers\textsuperscript{[67]}) at a surface, causing a change to the resonance energy of a propagating light-electron excitation (plasmon-polariton) at the metal-analyte solution interface.\textsuperscript{[68]} SPR/GCW requires an intricate setup and is highly assay-specific and thus unsuitable for unknown sample determination. Colorimetry/spectrophotometry provides a visual test based on the absorbance of light, however, the technique may be hampered by a subjective analysis and limited in terms of specificity, detecting only certain classes of compounds and therefore, it often requires further verification by more sensitive and specific laboratory-based techniques.\textsuperscript{[47,69]} Speed and reagent costs are additional limitations of colorimetric methods.\textsuperscript{[64]} Spectrophotometry has been used with saliva samples to determine the concentration of glucose, the abnormal concentration of which may be indicative of diabetes, requiring a 90-min
1.3.3. Raman spectroscopy

Raman spectroscopy (RS) is performed with a monochromatic light source, optics to remove unwanted light, and a spectrograph/monochromator to isolate a specific wavelength range. The technique has also benefited from the improved capabilities of cameras and advances in analysis software. The Raman scattering phenomenon relies on the instantaneous inelastic interaction of light with molecular vibrations, whereby a change in bond polarizability as a function of nuclear motions results in an alteration to the emitted frequency of light known as Raman-shift. Different molecules present different bonds to analyze and therefore, differences in Raman peak energies and intensities. Thus, Raman is often termed as a “molecular fingerprint.” RS has a long history of analytical uses including explosives detection, food technology, and even in the analysis of artwork. Berger may have been the first to suggest that Raman spectroscopy could be used to analyze biofluids in a near-infrared Raman study of blood, and many studies have followed. Recently, its potential to become a clinical tool for early disease diagnosis has been highlighted.

The phenomenon of Raman scattering produces inherently sharp spectral peaks, unlike fluorescence spectroscopy, and thus Raman facilitates multiplexed measurement, which is beneficial for sensing purposes by allowing maximal information extraction at a minimal time and cost. The accurate determination of various diseases such as cancer or traumatic brain injury (TBI) is often dependent on the detection of multiple biomarkers. For instance, in the case of TBIs, identifying a suite of biomarkers may be necessary to differentiate between demyelinating disease, polytrauma, or a co-morbidity that otherwise affects blood-brain barrier integrity. Rehman and coworkers have tabulated and assigned Raman peaks from the literature across a range of biological tissues.

RS further confers the advantage of requiring small volumes of samples (μLs), which is pertinent when using saliva as it is challenging to acquire rapidly in large volumes. Significantly, Raman scattering does not suffer from interference from water molecules (99% salivary constituent), as does infrared absorption spectroscopy, due to the low Raman cross-section of water. RS does not require sample staining, again unlike fluorescence-based analysis, and therefore, has been widely exploited for studies of living cells, unperturbed, in their native environment. Raman has several notable setup variations, including integration with interferometry (Fourier-transform Raman) and confocal microscopy (Raman micro-spectroscopy). Other variations are phenomenological. Coherent anti-Stokes Raman scattering (CARS) is a non-linear optical analog that can provide extra sensitivity. Resonance Raman spectroscopy (RRS) relies on exciting vibrational bonds at a laser wavelength close to resonance for an increased signal and has been employed in several studies incorporating Raman and saliva. Most notable, however, is surface-enhanced Raman spectroscopy, where Raman scattering is combined with plasmonic materials, supporting electron-light
excitations at a surface to considerably increase the Raman signal, often by many orders of magnitude.\cite{12}

**1.3.4. Surface-enhanced Raman spectroscopy**

First reported by Fleischmann, McQuillan, and Hendra in 1973/74\cite{109,110} while studying pyridine at roughened silver electrodes, surface-enhanced Raman spectroscopy (SERS) is a technique that can be chiefly understood in terms of large electric fields generated by surface-confined, hybridized electron-light excitations (plasmon-polaritons) associated with metals, usually gold or silver, at the nanoscale.\cite{68,87,111} These large local electric fields, termed “hot-spots” when concentrated to a small gap, couple to photons participating in Raman scattering events, leading to significant Raman signal enhancements. The enhancing factors can reach a $10^8$ increase over a non-SERS regime for a substrate-averaged measurement,\cite{111} or more if nanometric substrate locations are isolated and evaluated. Regions where the electric field is most concentrated can disproportionately affect the observed SERS enhancement.\cite{112,113} Concurrent with the electromagnetic SERS effects, “chemical enhancement,” consisting of alterations to bond polarizability upon molecule surface adsorption, is also broadly discussed in the SERS literature, however, the magnitude and extent of impact remain a matter of debate.\cite{14,87,111,114} Alessandri and Lombardi\cite{115} have recently reviewed non-electromagnetic effects in SERS in dielectrics. Similar to ordinary Raman, SERS also depends on the inherent cross-section of the analytes as well as the number of molecules present and their orientation on the enhancing surface.\cite{81,87,106}

**1.3.5. Potential of Raman and saliva, and current state of play**

Raman scattering is inherently selective, leading to the potential of accurate determinations, meanwhile, SERS can provide increased sensitivity and low limits of detection (LoDs). Saliva has comparable diagnostic potential to other biofluids. The ease of acquisition of saliva coupled with the speed and portability of RS can facilitate continual monitoring, crucial where the nature of a medical emergency is time-sensitive or temporal kinetics are required, such as in post TBIs.\cite{49,116} Thus, the combination of Raman spectroscopy and saliva is attractive for translation to the clinic and portable uses at the point-of-need. Deriu et al. show that a Raman approach for cannabinoid detection in saliva is almost three times faster than ELISA, despite involving a SERS preparation step (36 min versus 120 min for synthetic cannabinoid) (Figure 2C),\cite{48,117} and considerably longer time-to-results have been shown with ELISA processes taking several hours.\cite{118} For illicit compound detection, while the time for blood analysis using conventional methods can take days to months, saliva analysis takes only 2–48 hours,\cite{11} which could be almost instantaneous with a portable Raman system and an established substance database.\cite{119}

In a recent review detailing nanophotonic approaches to pharmaceutical monitoring via Raman, Frosch et al.\cite{120} mention saliva as a potential breakthrough diagnostic biofluid. Previously, Butler et al.\cite{12} set out a protocol for Raman studies with biological materials, and Henson and Wong\cite{121} outlined the optimal procedures in the collection, storage and the processing of saliva samples in the context of oral biology, while
Chevalier et al.\cite{122} provided recommendations on storage in a detailed study into proteomic longevity. Despite this literature, there is much room for researchers to take different approaches in setting up their Raman-saliva experiments and the methods they follow, and therefore, a summary of these details would be helpful. Importantly, there are other aspects of saliva analysis, specifically relating to the use of RS, such as the exact measurement protocol, that have not been adequately surveyed.

2. Applications of Raman-based saliva studies

Applications of saliva have been summarized previously with impact in the fields of specific disease identification,\cite{123} illicit drugs\cite{124} and pharmaceuticals.\cite{120} We briefly summarize the main applications of Raman-saliva study here. An overview is provided in Figure 3A.

2.1. Healthcare

Healthcare is the most significant potential application, where saliva is increasingly a possible candidate for auxiliary diagnosis.\cite{125} Saliva exhibits potential for use in the developing world, where the characteristics of diseases are poorly defined and treatment options are often limited, unavailable, or ineffective.\cite{20} Sialometry (salivary flow changes) and sialochemistry (salivary chemical changes) have been used to monitor general health,\cite{126–129} and these could be easily combined with RS.

In certain areas of medicine, saliva is considered a possible first-line diagnostic supplementing existing diagnostic processes.\cite{123} For instance, in an ovarian cancer study, Zermeño-Nava et al.\cite{125} note that sialic acid detection in saliva via RS should be “considered in the clinical scenario of adnexal mass growth diagnosed patient and not in the population in general” as well as combined with further clinical diagnosis (Figure 2D). This may be a good indication of the ambit of saliva in healthcare studies in general, whereby aid may be given in confirmatory diagnosis, risk stratification, prognosis determination, and therapy response monitoring\cite{20} or preventative screening.\cite{130}

Although the application to population monitoring is apparent,\cite{20} the complex and variable range of saliva constituents\cite{53} combined with the low specificity of certain compounds\cite{125} may limit the ability of saliva as a diagnostic biofluid in the determination of unknown disease, i.e., no suspected pathology.\cite{9} Most healthcare saliva studies use human samples from diseased subjects. Radzol and colleagues have used saliva spiked with nonstructural protein 1 in initial chemometric-focused SERS studies into Dengue fever.\cite{131,132}
2.2. Illicit drugs

Illicit compounds have been studied extensively with Raman spectroscopy, and saliva has become an increasingly popular biofluid for these studies. The majority of the early work in studying drugs in saliva has been led by Farquharson et al. The practicalities of developing a SERS sensor to detect illicit drugs, including via saliva, have been recently reviewed by Yu et al. The authors note that conveniently many illicit compounds are also good Raman scatterers. Raman-saliva studies in illicit compounds show high sensitivity using SERS-based sensing, comfortably outperforming the 10 s ng/mL range cutoff for many illicit drugs as recommended by the US Substance Abuse and Mental Health Services Administration (SAMHSA).

Sivashanmugan et al. have analyzed saliva in a Raman study on cannabis users, acquiring samples 15 min after the established use. Similarly, in a study on methamphetamine, Qu et al. successfully acquire 20 saliva samples for SERS analysis from actual methamphetamine addicts from the Residential Drug Treatment Center of
Beijing You-An Hospital, reporting discrimination between these and the saliva analysis of 20 non-addicted subjects. Collection from bona fide drug users is useful, but unusual, for reasons of legality, ethics, and compliance, and likely to be even more difficult to conduct for more harmful compounds i.e., heroin, cocaine. Similar methamphetamine studies have used spiked saliva samples.\cite{47,140}

2.3. Forensics

Saliva has been identified as a potential medium for forensic analysis via RS.\cite{149} Lednev and coworkers have pioneered extensive studies into the use of RS with bio-fluids for forensics, including saliva.\cite{2,102,150–154} The authors have conducted studies to identify different phenotypes including sex,\cite{155} and have differentiated between human and animal samples.\cite{156} Recently, Buchan et al.\cite{157} have also discriminated between male and female subjects, at differing age groups, via a self-organizing map clustering algorithm displaying a sex classification accuracy of 93%. Virkler and Lednev\cite{152} noted

---

Figure 3. Methods in Raman and saliva. A. Analysis of smokers’ saliva with Raman. (i) Raman spectra of S-S stretch and amide I derived from disulfide bonds of saliva proteins. The black line and gray line correspond to stimulated saliva and unstimulated (resting) saliva, respectively. (ii, iii) The areal ratio of disulfide residues per amide I for (ii) smokers, and (iii) nonsmokers. (iv) Sulphhydryl residues in the saliva of smokers and nonsmokers collected under stimulated and unstimulated (resting) conditions were compared separately with or without L-ascorbic acid 2-phosphate (Asc-2P). Asterisk represents significant difference ($p<0.05$). Adapted with permission from Taniguchi (2013) under Creative Commons BY 3.0. © Hindawi 2013. B. SERS detection of drugs in saliva with magnetic nanoparticles. (i) SERS spectra of pure saliva recorded (a) without and (b) with Au-dotted magnetic nanocomposites (AMN), (c) SERS spectrum of cotinine (5 $\mu$M) in saliva with AMN, (d) magnetically optimized SERS spectrum of cotinine (5 $\mu$M) with AMN. (e; Inset) High-resolution TEM image of AMN. (ii) Magnetically optimized SERS spectra of cotinine with various concentrations in saliva, mixed with AMN. (iii) (a) SERS spectrum of pure saliva recorded with AMN; (b) SERS spectrum of benzoylecgonine (50 $\mu$M) in saliva with AMN; the magnetically optimized SERS spectra of benzoylcegonine with different concentrations of (c) 0.1 $\mu$M, (d) 0.5 $\mu$M, (e) 5 $\mu$M and (f) 50 $\mu$M in saliva, mixed with AMN. Adapted with permission from Yang (2015). © Elsevier 2015. C. Raman and saliva for identification of novel synthetic cannabinoids. (i) Density Functional Theory simulated Raman (bottom, blue) and solid state experimental normal Raman (top, red) spectra of synthetic cannabinoid JWH-018. Intensities have been scaled for ease of comparison. (ii) Normal Raman spectrum of JWH-018 at a concentration of 5.0 mg/mL in methanol. Asterisks indicate bands assigned to the solvent. The intensities have been normalized to the solvent band at 3200 cm$^{-1}$ assigned to H-bound ν(OH) of the alcohol. (iii) SERS spectra obtained from fortified oral fluid extracts at different initial concentration of JWH-018. The SERS spectrum of the matrix blank is reported in gray. The vertical dotted line marks the drug’s reference band at 1370 cm$^{-1}$ assigned to the ν(C=O) of the naphthalene moiety. Adapted with permission from Deriu (2019). © American Chemical Society 2019. D. Analysis of sialic acid in saliva with Raman in ovarian cancer. (i) Cloud plot of sialic acid (SA) concentration for groups of benign and malignant cancerous patients. Insert: Box plot of sialic acid concentration. Red: ovarian cancer (OC); black: benign adnexal mass (BAM). (ii) Receiver operating characteristic (ROC) curve analysis for SA (and competitor biomarker cancer antigen 125: CA-125), to optimize the threshold of sialic acid concentrations to distinguish between BAM and OC patients. (iii) Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for each test. The figure includes cutoff values and confidence intervals for each biomarker, as well as for the calculated Malignancy Risk Index (MRI). CA125 > 35, CA-125 larger than 35 units/mL; CA125 > 216.5, CA-125 larger than 216.5 units/mL. Adapted with permission from Zermeño-Nava (2018) under Creative Commons BY 4.0. © Springer Nature 2018.
that the spectroscopic signature of saliva could be discriminated from those of blood and semen, and further that saliva samples from multiple donors were similar. Elsewhere, the same author has assigned the most relevant Raman peaks for blood, sweat, saliva, semen and vaginal fluid. Zapata et al. has reviewed the potential of spectroscopy for forensic biofluids.

2.4. Dentistry and orthodontics

Raman and saliva in combination are often used in dentistry and orthodontics, although saliva is viewed as a storage medium in most of these studies, focusing on spectroscopic analyses without the direct measurement of artificial saliva, instead using it as a simulated storage environment. Gurchukov et al. have studied periodontitis with Raman and actual human saliva samples, as have others in the contexts of remineralization.

2.5. General studies

Mleczko et al. study the interaction of antibodies and antigens in saliva and the effect of magnetic hyperthermia, mediated by hematite (Fe₃O₄) NPs acting to change local temperature and pH. Karlinsey et al. has studied the nucleation phase of hydroxyapatite on metal oxide with Raman spectroscopy in saliva, which may be of broad interest in biocompatibility studies with possible dental applications. In these studies, saliva is not interrogated by RS directly, and thus, like many of the dentistry studies, is less relevant to the current review. SERS studies of Yuen et al. into optimizing a gold-covered bead substrate have a “SERS substrate development character” and are not tethered to any one application. Other investigations also have a primary SERS substrate optimization characteristic.

2.6. Viral strains

In light of the COVID-19 pandemic, we note a recent study where Eom et al. has reported the use of SERS for the detection of mutant influenza in saliva and nasal fluid samples with spiky gold nanoparticles (AuNPs) and simple aptamer-functionalized glass slide substrates with 250 times greater binding affinity for the mutant pH1N1/H275Y influenza virus than for the wild-type virus. The authors note that although current diagnostic approaches can identify viral subtypes, they do not indicate antiviral drug-resistant strains. Recent publications have discussed the role of saliva specifically in identifying the novel coronavirus and the potential role of vibrational spectroscopy in COVID-19 identification.

3. Collection of saliva

3.1. Means of salivary collection

Typically, “whole saliva” is collected. This term refers to oral fluid from the salivary glands as well as from the gingival fold, oral mucosa transudate, nasal cavity, and
pharynx regions. The unprocessed mixture contains not only a plethora of proteins but also nasal and pharyngeal mucus, micro-organisms, desquamated epithelial cells, and blood cells, as well as large pieces of food debris. There may be significant variations in the degree of salivary interference from exogenous stimuli between different subjects.

Unstimulated salivary flow is the basal flow at rest, whereas stimulated flow is induced by mechanical, olfactory, gustatory or pharmacological stimuli. Saliva may be collected by different means, including, for instance, swabbing or suction (Table 1). The most common approaches for unstimulated saliva collection are the passive drool method or simple spitting, which require no specialized training and are noninvasive. Stimulated salivary flow may be produced by supplying the subject with a piece of paraffin to chew on or by placing a drop of citric acid on their tongue. Machado et al. employed dental gauze rolls which study subjects kept in their mouth for 3 min before centrifugation was applied to extract the saliva, a method earlier described by Chiappin. Inscore et al. employed foam-head swabbing coupled with syringed extraction to collect saliva samples.

### 3.2. Volume of the collected saliva

Saliva volumes acquired are typically in the mL range, which is often more than sufficient for Raman analysis, even via microfluidic Raman systems, and acquired easily, unless due to specific disease hyposalivation or the subject’s severe dehydration in sport and exercise studies. Indeed, individual hydration is the most crucial factor in salivary flow, and when bodily water content reduces by only 8%, the salivary flow is effectively zero. Barring dehydration, the amount of unstimulated saliva generated per minute in a healthy adult is 0.25–0.35 mL, rising to up to 3 mL/min when stimulation occurs. In a study by Taniguichi et al., the subjects were required to provide saliva via draining/drooling for 7.5 min or until 20 mL was acquired. Exercise may also cause the salivary composition to exhibit increased levels of α-amylase and electrolytes, in particular sodium, depending on the intensity of the activity. However, it is not clear within most Raman studies how saliva may be different for participants who arrived at the point of sample deposition by different levels of exertion i.e., a range of more modest activity levels.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting</td>
<td>&lt;1 hour</td>
<td>[47,96,103,213,214]</td>
</tr>
<tr>
<td></td>
<td>1–6 hours</td>
<td>[83,131,140,189,190,193,207,215,216]</td>
</tr>
<tr>
<td></td>
<td>Overnight</td>
<td>[98,202,217–220]</td>
</tr>
<tr>
<td>Mouth Wash (water)</td>
<td>Immediately prior</td>
<td>[97,98,202,217,218,220–227]</td>
</tr>
<tr>
<td></td>
<td>&lt;1 hour</td>
<td>[60,96,99,100,193,213,215,228]</td>
</tr>
<tr>
<td>Mouth Wash (other)</td>
<td>Alcohol-free</td>
<td>[125,200,201,229]</td>
</tr>
<tr>
<td></td>
<td>Stimulated</td>
<td>[69,146,189,207,232,233]</td>
</tr>
<tr>
<td>Method</td>
<td>Collection container (i.e., spitting, passive drool etc.)</td>
<td>[50,137,146,232,234–235]</td>
</tr>
<tr>
<td>Swab</td>
<td></td>
<td>[69,138,233]</td>
</tr>
<tr>
<td>Lashley cup/suction</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Age, menopause, and hereditary factors also play a role in the flow rate of saliva. With age, salivary glycoproteins increase as a consequence of innate immunity. In other cases, changes may be less due to aging and more due to ill-health and certain medicines used in older subjects, such as anticholinergics (neurotransmitter blockers), which may induce hyposalivation. If the subject is ill, the overproduction of mucin may be a problem and sample dilution may be necessary prior to analysis.

3.3. Collection in Raman-saliva studies

Recently, Goodacre has emphasized that sampling procedures should be considered an essential aspect of the analysis of complex natural systems, and notably, Taniguchi et al. conveyed that differences in diagnostic performance depended on whether saliva collected was stimulated or unstimulated in a study of mucin in smokers’ saliva (Figure 2A). Many collection protocols in the Raman-saliva literature are strict, albeit inter-study differences are still significant. For example, in a study of malignancy in breast tissue, Feng et al. incorporated a 12-hour fast, a narrow collection window (6:30–8:30am), and three mouthwashes, demonstrating a statistically significant difference in SERS peak intensities \( p < 0.05 \) between healthy and cancerous breast tissue samples. Lin et al. subsequently employed almost identical measures in a nasopharyngeal carcinoma study within a microfluidic device.

In dentistry, Axelsson has indicated that while fasting reduces salivary flow, it does not lead to hyposalivation. Maitra et al. extends abstention to liquids, however, with only a three-hour fast before collection. Similarly, Taniguichi et al. dictates no cigarette usage for at least three hours prior to saliva acquisition in their study of salivary mucin changes in smokers. Dietary and lifestyle aspects of participants, such as BMI or dental hygiene, are generally not recorded across studies. Malkovskiy et al. note no apparent dietary influence, with just a diurnal variation in salivary thiocyanate concentration present in a study into cystic fibrosis with RS (Figure 4B). Salemmlani et al. impose no prescription medicine to be taken prior to sample collection. This highlights an important factor since hospital patients often provide samples early in the morning, many of which may have co-morbidities.

Kah et al. required the study subjects to wash their mouth 30 min before collection and refrain from swallowing for several minutes to aid collection. It is unknown whether this impacts salivary composition. Radzol et al. ask volunteers to perform 1 min of gargling before unstimulated collection. Hernández-Arteaga et al. implemented “vigorous teeth brushing,” an approach which could introduce unwanted blood residues into the saliva, which is in the exclusion criteria of Othman et al. Salivary pH upon collection, in general, is not considered, barring ref. the spectral impact of which has been discussed in Buchan. Establishing a universal “standard collection time” will be helpful in the consistency and comparison of the many studies exploiting saliva, as has been noted in the context of CSF and serum acquisition.

In a lung cancer study, Li et al. notes that similar numbers of smokers in their control group \( n = 13/21, 65\% \) were selected in their lung cancer cohort \( n = 14/20, 67\% \). This is a sensible choice given the effect that the smoking phenotype,
Independent of the cancer status, can have on the salivary Raman spectra.\textsuperscript{146,207,216,239–242} The authors perform a baseline check by ELISA for possible inherent morphine traces in the hospitalized cohort. Hernández-Arteaga\textsuperscript{200} et al. further classify patients as with “no systemic disease” other than the disease being studied, i.e., breast cancer, as well as with “no oral complaints” due to the nonspecificity of sialic
acid, which is elevated in other cancers and inflammatory conditions. However, such restrictive measures may prove problematic in some point-of-care settings.\cite{7}

Less stringent approaches for the saliva collection criteria include the direct collection of unprocessed saliva from healthy individuals without any prerequisites.\cite{243,244} While this might introduce more significant intra-class variability in research studies, it is more in line with real-world applications, at the pitch-side, roadside, or at home, with results only later relayed to clinicians. Such applications need to be well-supported with technological advancements including AI and machine learning algorithms as decision support tools.\cite{70,245–247} It is plausible that common interferents such as mouthwash or alcohol have less of an effect on substance identification than expected, as evidenced in illicit compound detection via infrared absorption spectroscopy.\cite{248}

### 3.3.1. Sample numbers

Raman-saliva studies in the literature involve both small sample numbers ($n < 20$)\cite{60,96,118,153,154,224,241} as well as larger ($n > 150$) cohorts (Figure 3B).\cite{200–202,222,238} Regarding small sample sizes of saliva, Kah et al.,\cite{60} who study five patients with oral cancer and five healthy subjects, indicate that a larger cohort is needed for statistically significant cancer staging. On the other hand, Maitra et al.\cite{83} collected nearly 500 samples for a biofluid cross-comparison study, including 114 saliva samples, further divided into 35 healthy and 79 diseased cohorts, spread across different stages of esophageal cancer. It is important to consider sufficient sample numbers in saliva studies due to relatively large variations in the salivary matrix constitution i.e., intra-class variation. This is especially relevant if the clinical stage of disease is being evaluated and the disease progression is associated with subtle changes in concentration of specific biomarker(s) such as, for instance, in the progression of early hyperplasia to invasive carcinoma in oral cancer.\cite{60,249} It is critical in certain pathologies where the survival rate drops significantly with advanced stages of the disease.\cite{200,201} Hernández-Arteaga et al.\cite{200} note a statistically significant difference between sialic acid concentrations in saliva interrogated by SERS in breast cancer stage 0 and 1 versus stage 3 and 4 cohorts. In this study, concentration relative standard deviation (RSD) amongst all pathological samples was 50%. For future clinical validation studies, power calculations will ensure sufficient study sample sizes to allow the determination of a clinically significant difference.\cite{250} Large sample numbers, however, are not necessary for spiked saliva studies, which may be produced by adding the required analyte to artificial/simulant saliva samples.

### 3.3.2. Saliva from commercial vendors

Saliva for Raman studies is frequently acquired from a commercial source. Purchased saliva sample studies include Shende (Lee Biosolutions),\cite{141} Muro (Lee Biosolutions and Bioreclamation),\cite{154} Muro (Biological Specialty Company, Lee Biosolutions, Bioreclamation),\cite{155} D’Elia (Bioreclamation),\cite{105} Eom (Lee Biosolutions),\cite{197} and Al-Hetani (Bioreclamation).\cite{239} Obtaining saliva from a commercial source may simplify subsequent steps in saliva storage and measurement and negate problems regarding collection and questions on optimal pre-processes e.g., centrifugation.
4. Pretreatment

Saliva is arguably easier to process than blood, which is prone to clotting and requires specific sample containers. RS also requires fewer reagents than other quantification techniques. There is, however, no universally set protocol for saliva pre-processing. In healthcare studies with human samples, centrifugation of the saliva matrix to remove debris and larger constituents is standard, albeit with differences in process time and centrifugation speed (Table 2). Analogously, Salemmilani et al. filter saliva samples using a 0.2 μm-diameter syringe to remove large cells and debris. Ma et al. further note that standard procedures of sample pretreatment and purification should be established to obtain more reliable and specific SERS spectra of biological species. The extent of sample pretreatment, or "sample enrichment," may depend on the concentration range of the biomarker present in the salivary matrix and the sensitivity required, as well as what may be practical. Special care may be needed where specific parts of the salivary matrix need to be isolated, for example, specific proteins, lipids (Folch method), or extracellular vesicles (EVs).

Owing to their viscous nature, mucins, which are large, glycosylated proteins, can trap substances of interest and potentially interfere with an interspersed plasmonic medium in SERS studies. This may mandate further pre-processing steps. With the addition of a solid-phase extraction step, Inscore et al. managed to measure half the LoD of cocaine in saliva compared to corresponding concentrations in water. Even with the inclusion of a subsequent nanoparticle-SERS step, the complete process was completed in under 10 min. Viscosity is also a problem for the complementary technique of spectrophotometry, which is suitable for elemental analyses such as calcium and magnesium concentrations in saliva. In such cases, saliva samples must be diluted, which is detrimental to analytical sensitivity. In an investigation into the detection of thiocyanate in smokers versus nonsmokers by Wu et al., saliva samples are diluted by a factor of 10 prior to 30 min 7000 g centrifugation (Figure 5). Alternatively, there may also be a need to perform such dilution to cause a reduction in the sample viscosity and thus ease of flow through a microfluidic channel.

D'Elia et al., while studying saliva and cocaine, avoid any pretreatment steps by transferring saliva to NMR tubes directly before resonant Raman analysis, an approach

Table 2. Pretreatment routines and storage of saliva as listed in the literature. Excludes drug mixture and artificial saliva use.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centrifuge speed (rpm)</td>
<td>&lt;10,000</td>
<td>[48,96,125,196,200,203,219,222,232,241,244,254,255]</td>
</tr>
<tr>
<td></td>
<td>10,000</td>
<td>[97,103,220,221,223,227,236]</td>
</tr>
<tr>
<td></td>
<td>&gt;10,000</td>
<td>[60,98–100,131,193,202,204,217,218,222,225,257,258]</td>
</tr>
<tr>
<td>Centrifuge time</td>
<td>1–2 mins</td>
<td>[96,232,256]</td>
</tr>
<tr>
<td></td>
<td>5 mins</td>
<td>[60,193]</td>
</tr>
<tr>
<td></td>
<td>10 mins</td>
<td>[48,97–100,131,202,204,217–221,232,224,227,254]</td>
</tr>
<tr>
<td></td>
<td>10–40 mins</td>
<td>[125,196,200,203,222,241,244,255,257]</td>
</tr>
<tr>
<td>Filtering</td>
<td>Methanol</td>
<td>[48,201]</td>
</tr>
<tr>
<td></td>
<td>Acetonitrile</td>
<td>[244]</td>
</tr>
<tr>
<td></td>
<td>Acetic acid</td>
<td>[137]</td>
</tr>
<tr>
<td></td>
<td>Chloroform</td>
<td>[256]</td>
</tr>
<tr>
<td></td>
<td>Syringe filter</td>
<td>[47] (0.2 μm diameter)</td>
</tr>
<tr>
<td></td>
<td>Membrane</td>
<td>[241] (0.45 μm pore diameter) [241] (0.22 μm pore diameter)</td>
</tr>
<tr>
<td>Storage temp</td>
<td>4 °C</td>
<td>[48,98,107,188,196,203,213–216,219]</td>
</tr>
<tr>
<td></td>
<td>–20 °C</td>
<td>[60,197,217,230]</td>
</tr>
<tr>
<td></td>
<td>–80 °C</td>
<td>[83,96,103,189,202,221–224,227,232,239]</td>
</tr>
<tr>
<td></td>
<td>Liquid Nitrogen</td>
<td>[97,220]</td>
</tr>
</tbody>
</table>
which may be more applicable for real-world portable sensors where point-of-need saliva samples cannot be easily or rapidly processed. This approach, however, significantly compromises signal intensity and the resultant LoD of 10 \( \mu \)g/mL is three orders of magnitude above the requisite forensic specification. In a recent study by Hole et al., the authors use RS to examine the impact of three modes of saliva pretreatment: air-drying,
lyophilization, and centrifugation (pellet and supernatant). They conclude that air-dried and lyophilized samples maintain the spectroscopic character of whole saliva, suggesting that minimal processing and freeze-dried storage are most suitable for discriminating pathological salivary samples.

5. Storage

Chevalier et al. have analyzed the longevity of salivary proteins via electrophoresis as a function of time since sample collection, storage temperature, presence or absence of a protease inhibitor, and the removal of insoluble materials. It is well-known that amylase in saliva can degrade salivary proteins. Shorter storage times, lower storage temperatures, with the addition of the enzyme inhibitor, and removal of large material, are all concluded to be beneficial. In alignment with the previous mass spectrometry report of Schipper et al., the authors conclude, “In case of a clinical comparison with a pathological condition, control saliva samples should be collected from a healthy nonsmoking subject, in the morning, at least 2 h after eating, and the mouth should be rinsed with water. After collection, saliva samples should be stored in a freezer at −20°C, and during sampling, saliva should be kept on ice with a protease inhibitor cocktail and centrifuged to remove insoluble material and then stored at −80°C.” Buchan et al. have recently reported minimal spectral changes in saliva analysis, with RS, over a seven day period when stored at room temperature.

In Raman-saliva studies, Inscore et al. measured drug-doped artificial saliva samples within 60 min of preparation, and Machado et al. has used saliva samples within one day of collection. Elsewhere, Maitra et al. have stored freshly acquired samples at 4°C–7°C before transfer to −80°C in a practical protocol. Many reports, however, do not provide storage details, implying that measurements were taken soon after obtaining the samples. While critical for lab-based work with large sample sizes or protracted studies, storage concerns may be considered mostly irrelevant for the instantaneous portable Raman systems using saliva on-site unless confirmatory steps are required in a laboratory setting at a later stage. Storage considerations might be further highly dependent on the entities desired to be preserved. For instance, exosomes, a class of EVs, are remarkably stable, and in fact, there is a greater concern with the integrity of any attached surface proteins that may not persevere well in ambient or high-temperature storage conditions.

6. Measurement protocols

Spectroscopic saliva measurements lack a standard protocol and thus issues with variability and irreproducibility persist. Maitra et al. and Muro et al. acquire 25-point spectra in saliva measurements, and Virkler et al., 36 random points across a 75 μm x 75 μm area. However, many studies appear to rely on considerably fewer measurements, as little as three. This can create further challenges for subsequent chemometric data analyses where only subtle differences are present in Raman peak intensities between saliva samples and intra-variance in any single sample class is high. Notably, in a biofluid forensics study, Lednev and coworkers, based on their prior...
individual biofluid investigations, tailor the number of Raman map points and integration times depending on the specific biofluid under interrogation, blood, sweat, saliva, semen and vaginal fluid.\cite{154} Elsewhere, Cottat et al.\cite{118} notes inhomogeneity in terms of the density of the affinity molecules for a liver cancer biomarker and thus acquires 10 separate point measurements across the surface. Wu et al.\cite{241} acquire Raman measurements at multiple different positions on a microfluidic chip, which may serve as a suitable protocol for quantitative measurements but is clearly an added complexity compared to automated single-spot measurements (Figure 5). Measurements on aluminum or titanium foil for suppressed Raman background signal are common in Raman and saliva studies.\cite{69,83,255,261}

6.1. Drop-casting and dried samples

A widely used approach for preparing liquid samples for Raman measurements is application by (micro)pipette and allowing to air-dry on a supporting surface.\cite{12,106,107,131,262,263} This is known as the “drop-casting” method and is viewed as a quick and easy approach within Raman studies.\cite{264} The technique often results in a highly non-uniform distribution of fluid constituents across the dried spot area due to the evaporation gradient and consequent capillary forces acting on the drying droplet.\cite{265} This somewhat overlooked problem\cite{12} could be mitigated via an absorbent substrate\cite{202,266} or by using the whole saliva fluid and an interspersed plasmonic medium to enhance the Raman signal. The effect can also be pertinent in a functionalized gold nanoparticle film, where the nanoparticles are drop-cast onto the surface before the saliva deposition. If accurate quantitative analysis is required, the exact regions being measured will need to be carefully identified.\cite{60,262}

In some cases, the “coffee rings,” or more precisely, radial surface distributions, can be beneficial,\cite{12,262,267–270} having been used for facilitating the detection of various biomolecules,\cite{263,271–275} and used for concentrating or spatially separating the desired salivary constituents.\cite{276,277} In a proteomic study, Zhang et al.\cite{270} showed that such ring formations remain stable for weeks. We note, the \( \mu \)m- and mm-scale structural patterns in dried drop-casted bio-samples can be used as a visual diagnostic tool,\cite{278} a “Litos test,” as noted by Sefiane,\cite{279} and initially found use in urine sample analysis for urolithiasis.\cite{280} Similarly, Gonchukov et al.\cite{188} noted irregularities of dendritic structure in dried periodontitis saliva samples prior to Raman analysis, and more recently, the technique has been deployed in conjunction with machine learning for discrimination of blood samples of healthy participants pre- and post-exercise with an accuracy of 95%.\cite{281} However, more subtle biochemical changes require the specificity of spectroscopic analysis for unambiguous detection.

In a study on vacuum-dried saliva samples, Malkovskiy et al. report that the location on the dried sample spot has no impact on the measurement. This may be due to the rapid vacuum-dried stage employed by the authors, which mitigates the acting time for capillary forces on the material in the drying saliva drop, resulting in a more homogeneous material distribution.\cite{69,265} Falamas et al.\cite{224} employ a lyophilization (freeze-drying) process. No apparent effect on the classification accuracy by varying measurement location has also been reported by Maitra,\cite{83} who have studied air-dried biofluids on
aluminum or titanium slides prior to measurement. Others conduct measurements on
dried saliva when investigating methamphetamine or sialic acid in saliva via Raman,
respectively.[99,100] In developing a paper-based substrate for SERS with a concurrent
ambient pressure mass spectrometry analysis, Díaz-Liñán et al.[282] dry analyte-spiked
saliva on a surface before swabbing the tip of the paper substrate across the dried saliva
area. While this protocol of swabbed sampling produces a less linear calibration curve,
and may induce damage to the sample, the inhomogeneity problems that arise from
conventional drop-casting can be mitigated.

Qian et al.[99] show useful optical images of the dried saliva (1 μL) drops, displaying
fern-like dendritic formations on the surface, the formation of which was discussed by
Pearce and Tomlinson.[283] in the context of tear studies. Such surface inhomogeneity
suggests the need for a larger number of measurements. Derjaguin-Landau-Verwey-
Overbeek (DLVO) theory, which describes forces between small particles in solution in
terms of electrostatic repulsion and van der Waals attraction,[284] can be used to predict
the interactions between small particles in solutions, and has been applied in conjunc-
tion with finite element modeling to predict the behavior of nanometric particles in a
drying drop,[285] although, the theory will need to be modified when applied to bio-
logical entities in a complex matrix.[286] Given the simplicity of the drop-casting
approach, the development of a saliva-specific model would be desirable.

A direct benefit of measuring air-dried samples is the ease of transportation that may
facilitate point-of-care detection and patients’ self-care remotely.[69] In conjunction with
developments in portable, easy-to-use lab-on-a-chip devices, this might serve as a cata-
ylist for decentralized medicine and a move away from a single disease diagnosis to a
more all-encompassing concept of “health surveillance” and monitoring.[130,287]

6.2. Alternative approaches

Measurement of the saliva in solution, i.e., in a native state, may result in the problem
where too few of the target molecules are within the illuminated laser area and conse-
quently, the Raman signal is too low to be of practical use. With a planar nanostruc-
tured SERS surface, the target molecules may not be within adequate nanometric
proximity to the substrate to experience a sufficient plasmonic enhancement.[288]
Therefore, unless the target is present at a high concentration, saliva analysis in a native
liquid form requires SERS analysis with interspersed nanoparticles. To note, in complex
biological matrices where many different molecular species are present, such as in saliva,
and where different moieties may have highly diverse binding affinities to the metal
nanoparticles, preferential adsorption could mean the exclusion of the requis-
ite analyte.[65]

Another approach for analyzing saliva in a liquid form might be to employ optical
trapping, as has been used recently for red blood cell analysis,[289] relying on radiation
pressure from incident light onto a sample to spatially confine a desired salivary con-
stituent, and this could be used in conjunction with a SERS-active medium.[290]
Structured light i.e., bespoke polarization, phase, and amplitude, in optical trapping (in
3D: optical tweezers), has recently been reviewed by Yang,[291] including deployment of
optical tractor beams, in what could be a kind of nano-factory[292] for manipulation of
biofluid constituents on the microscale. Elsewhere, inexpensive and highly absorptive paper-based substrates, which can be simply dipped into the salivary medium, could offer fast and homogeneous detection.[266,293,294] Zangheri et al.[295] explore a paper-based chemiluminescence sensor for salivary cortisol, but this approach hitherto appears to be untested in Raman-saliva studies.

6.3. Raman-saliva studies with functionalized surfaces

Sensing platforms can be designed to be highly analyte-specific by utilizing a range of affinity molecules anchored to the detection surface to provide the required selectivity for the analytes of interest.[67] In some cases, the measured Raman signal is not from the analyte of interest but from a tag molecule in a “labeled” detection assay, where the measurement sensitivity is determined by proxy.[106] Raman peaks of the analyte and affinity molecules may overlap spectrally, and this can hinder analyte detection. Within Raman-saliva studies most investigations avoid any intermediary molecule (functionalized surface, affinity molecules, labeled detection) and use an unmodified Raman detection regime,[10] which is more cost-effective for real-world portable systems.

7. Raman measurement parameters

7.1. Wavelength

A majority of Raman-saliva studies employ a lab-based commercial system (Figure 3D),[12] most commonly applying an excitation laser wavelength of 514/532 nm, 633 nm or 785 nm (Figure 3C). As an exception, D’Elia et al.[105] uses stimulated resonance Raman at 239 nm excitation, establishing a LoD of 10 μg/mL for cocaine in artificially mixed saliva-cocaine samples while pointing out that a laser excitation at 200 nm could decrease the LoD further with significant sensitivity improvement required for forensic drug detection (8 ng/mL). The movement toward shorter excitation wavelengths confers a significant benefit in terms of signal due to the $1/(\text{wavelength})^4$ dependence of the Raman scattering intensity. Perhaps just as important, UV excitation avoids the need for any pre-concentration or other preparatory step by mitigating the inherent, and masking, effect of absorption from the native oral fluid. However, the well-known damage that UV irradiation may pose to mammalian cells means that the extension of use of UV excitation to healthcare studies might be tentative. Due to the typically large fluorescent background in the UV and visible wavelengths, 785 nm excitation dominates Raman-saliva studies. Clinical Raman applications necessitate the need to balance the inherent wavelength dependence of the Raman scattering cross-section, the presence of the fluorescent background, as well as the quantum efficiency (QE) of any detectors used[86] where the QE of silicon-based CCDs decreases precipitously in the near-infrared (NIR) spectral range.[296] In micro-spectroscopy applications, the wavelength will also affect the lateral resolution and the probed sample depth.[296,297]

Hernández-Arteaga et al.[200] use a green laser at 532 nm because 785 nm excitation was found to cause an evaporation of the salivary water medium, using silver as the plasmonic enhancing metal, which, unlike gold, exhibits no electronic inter-band transitions in the visible part of the spectrum below 600 nm.[14,68,298] Often, the plasmonic
properties of nanostructured SERS substrates dictate the chosen excitation wavelength or vice versa, such as in the study by Cottat et al. [118] who use a 660 nm laser to detect a biomarker for liver cancer via a nanocylinder surface and near-field coupled nanorods. Aluminum (in the nanostructured form) is a well-recognized alternative plasmonic material in SERS [82, 299, 300] but is also often used as a planar substrate in Raman studies to confer a low background signal rather than plasmonic enhancement [261].

Resonance of particular substances being investigated might further affect the selection of the excitation wavelength, should for instance, a narrow range of compounds or similar physiological bodies be the target. 1064 nm, for example, has become frequently used in plant studies [301] and has been used in mineralization analysis in dental investigations [302, 303] usually in the form of Fourier transform Raman, which permits fluorescence rejection and a better signal-to-noise profile (Fellgett advantage) [143, 304].

7.2. Laser power and optics

Raman-saliva studies often do not include details on the system optics or laser power at the sample surface, i.e., system losses between the source to sample. However, these parameters may play an important role in determining an accurate assessment of the likely laser photo-damage, and the signal uniformity of the acquired measurements since larger laser interrogation areas at the sample surface produce an inherently more uniform signal. When reported, it is clear that a broad range of objective lens types is used ranging from 10× [60, 125, 200, 201, 257] to 100× magnifications [96, 118, 244]. In portable Raman systems [99, 100, 146, 226], visible in both healthcare and illicit drug applications (Figure 3G), the traversal of the beam through the fiber optics may induce high losses. Zhang et al. discuss the stability of such systems, and Pence et al., the optimization of fiber-optics in Raman for clinical applications [46, 86].

When studying biological materials, minimizing laser exposure can be critical to the consistency of signal measurement. Possible photochemical (bleaching) or sample burning effects can be easily monitored with time-series spectral acquisition. Moreover, induced graphitification may lead to the appearance of artificial carbon D and G spectral bands. Thus, it is paramount to consider the exact substance being detected. For example, Farquharson et al. report single spectrum acquisition times of up to 300 s in the study of overdose drugs (Figure 6B), and D’Elia et al. use long exposure times of 30 s with 20 accumulations in the detection of cocaine in saliva [105, 138]. No ill effects are reported. Contrariwise, in a thiocyanate study, Malkovskiy et al. [69] note photobleaching in saliva samples when prolonged high laser power is used, specifically the appearance of S-O Raman features in the 1000–1100 cm⁻¹ region. In SERS measurements, plasmonically driven chemical or thermal effects may further need to be considered [87].

While surface powers of ≤ 5 mW are sufficient for laser interrogation in the green region [39, 257] and even for 633 nm laser excitation [47, 60, 234], higher laser power (≥ 10 mW) is often necessary in the infrared range [50, 145, 152, 154–156, 304] due to the 1/ (wavelength)⁴ Raman scattering signal dependence. The exact power density relies not only on the laser power at the surface but also on the numerical aperture (NA) of the objective lens and the excitation wavelength. Given the frequently used Raman parameters in the Raman-saliva studies, this value would appear to be, theoretically, in the
Figure 6. Numerical methods in Raman-saliva studies. A. Optimization of gold-coated polystyrene (Au-PS) bead SERS substrates for bioanalysis. SEM images of Au-PS substrate with different microwave irradiation times $T_{MW}$ (600 W, at 2.45 Hz): (i) 0 s, (ii) 200 s, and (iii) 600 s. (iv–vi) are the corresponding graphical diagrams with different geometries estimated from the SEM images with different $T_{MW}$ for discrete dipole approximation (DDA) modeling. In DDA modeling, the Au-coated PS bead sizes are assumed to have diameters (a) of 500 nm at $T_{MW} = 0$ s. The center-to-center distance (b) between adjacent PS beads is 500 nm at $T_{MW} = 200$ s. In the calculation, the geometry of PS beads is defined by physical contacts with adjacent beads over an edge length (c) of 176 nm positioned to another edge (d) at an angle of 120° with an arch (e) connecting the two ends of edges at the arch tangent point. At $T_{MW} = 600$ s, the beads change to hexagonal shapes with side (f) of 288.68 nm in the DDA model. The corresponding $|E|^2$ field distributions of the substrates with 785 nm light excitation under different $T_{MW}$ of 0 s, 200 s, and 600 s, are shown in (vii), (viii), and (ix), respectively. Adapted with permission from Yuen (2010). © John Wiley and Sons 2010. B. Silver nanoparticle electro-migration in dielectrophoretic microfluidic chip for SERS detection of drugs in saliva. Cross-sectional view of the microfluidic channel at midplane of one of the trap zones. Black lines correspond to contours of constant $|E|^2$. Red lines show pathlines of the nanoparticles toward the trap. Highly polarizable nanoparticles migrate from the regions of low electric field intensity to regions of high electric field intensity. Adapted with permission from Salemmilani (2018). © American Chemical Society 2018. C. Optoplasmonic SERS platform for detection of methamphetamine in biofluids. (i) Illustration of model representing optoplasmonic hybrids. E-field intensity maps with time-averaged Poynting vectors (cyan arrows) in the X-Z plane of FDTD-simulated models at the wavelength of 785 nm for (ii) AuNP monolayer, (iii) SiO$_2$ sphere with diameter = 500 nm, (iv) optoplasmonic unit with SiO$_2$ sphere diameter = 500 nm, (v) SiO$_2$ sphere with diameter = 2 μm, and (vi) optoplasmonic unit with SiO$_2$ sphere diameter = 2 μm. (vii) FDTD-simulated E-field map in the X-Z plane of AuNP monolayer and optoplasmonic units with different microsphere sizes: (viii) SiO$_2$ particle diameter = 500 nm and (ix) SiO$_2$ particle diameter = 5 μm. (x) Corresponding statistics on the percentage of different enhancement values in the E-field map of X-Z plane. All incident wavelengths in (vii–x) are 785 nm. (xi) SERS spectra of para-mercaptoaniline obtained on the optoplasmonic unit with varied diameter (500 nm to 6 μm) of the dielectric sphere. (xii) SERS sensing efficiency $\gamma$ and E-field enhancing efficiency $F_E$ as a function of dielectric particle diameter. The values are based on the SERS peak intensities at 1077 cm$^{-1}$ in (xii). Adapted with permission from Hong (2020). © American Chemical Society 2020.
range $10^5$–$10^7$W/cm² ($1$–$100$mW/µm²) in most studies. However, assuming an effective NA in excitation which could be two orders of magnitude lower, i.e., $\text{NA} = 0.01$ where the narrow, incident laser beam does not avail of the full width of the focusing objective lens, surface power could be $10$–$10^3$W/cm² ($1^{-4}$–$1^{-2}$mW/µm²). These values are rarely reported explicitly. Occasionally, a 3D volume is interrogated, obviating the need for any such calculation, as in Farquharson et al. Changes in blood samples due to the photoinduced protein denaturation, followed by hemoglobin aggregation, have been identified at higher laser powers. The effect of potential photodamage to the salivary medium specifically is yet to be established.

8. SERS studies of saliva

Almost three-quarters of the Raman-saliva studies surveyed include some form of SERS (Figure 3F) since many compounds in saliva, whether physiological or pharmaceutical in origin, are present in small quantities and thus often require an ultra-sensitive detection method. While early SERS studies centered on roughened electrodes, and nanoparticles remain in common use, nanostructured surfaces as enhancing media have been gaining interest. Intricately patterned top-down fabricated SERS substrates and inexpensive, high sensitivity, bottom-up SERS platforms have been emerging, building upon the significant advances in nano-fabrication technologies. Alternative approaches also exist, for instance, Su et al. use large AuNP clusters (Figure 6A), and recently, Velička et al. have used an electrochemical SERS silver electrode set-up to detect caffeine in saliva. SERS is viewed as a promising route to accelerate the adoption of Raman spectroscopy for biostudies. In a recent breast cancer study, Feng et al. note that healthcare studies using standard Raman spectroscopy might be inhibited by the small Raman cross-section of protein bands and large fluorescent background signal with low sensitivity to, often subtle, biochemical changes. Most studies within SERS employ gold or silver as the plasmonic medium. Gold is biocompatible being a highly inert material, while silver has well-known toxicity in bio-systems, and silver nanoparticles (AgNPs) have been shown to produce toxic effects.

SERS can now be viewed as an analytical technique, although the emphasis is on the end-user as to what reproducibility is required. It is widely accepted that depending on the application, there are various requirements from SERS in terms of the analytical sensitivity, signal uniformity, and reproducibility, with Bell et al. recently highlighting the methods to standardize SERS measurements, including better analyte control and instrumental factors. The non-linear nature of SERS entails increasingly large enhancements as the surface features become truly nanometric and the electric fields are increasingly localized. Notably, Fang et al. showed that 24% of the SERS signal originated from a mere 63 out of 1,000,000 surface sites on a silver nanoparticle SERS surface. However, there is also progressively less control of the morphology of the surface features as they get smaller, and this tradeoff between sensitivity and feature control is often termed the “SERS Uncertainty Principle.”

SERS has the potential to detect a wider range of compounds and provide better detection threshold than ELISA ($10^{-6}$–$10^{-8}$M) or high-performance liquid
chromatography (HPLC) fluorescence (10^{-7}M).\textsuperscript{118} Durucan et al.\textsuperscript{335} have introduced a SERS-assisted chromatography device using a nanopillar platform, which considerably improves sample throughput compared to mass spectrometry.

### 8.1. SERS sensitivity, reproducibility and reusability

In contrast to the trends in analytics,\textsuperscript{236} SERS studies routinely report figures of merit. However, there is a significant variation in the reporting of SERS substrate performance in the literature\textsuperscript{113} with differences in enhancement reference methods and enhancement factor calculations.\textsuperscript{14,106} Zhang et al.\textsuperscript{257} for instance, use bulk dye powder as the reference, while other researchers in the broader SERS literature report the use of solutions in cuvettes, and these vary further in how the liquid reference is calculated.\textsuperscript{113,317} Comprehensive details on a range of procedures and calculations for SERS are given by Le Ru and Etchegoin.\textsuperscript{14} The SERS enhancement factor (EF) is a measure of the increase in the Raman signal of a characterizing analyte molecule or the requisite compound in the study, compared to an unenhanced reference sample. Alternatively, more readily understood metrics such as the LoD and the Limit of Quantification (LoQ) can be used to quantify the sensitivity and performance of SERS substrates.\textsuperscript{14,332}

Reproducibility of SERS measurements is reported less often, despite having been discussed by Natan at the first Faraday Discussion on SERS in 2005.\textsuperscript{334} Wang et al. characterizes the reproducibility and uniformity of the signal performance of a tightly packed gold nanoparticle-on-glass substrate by analyzing the relative standard deviation (RSD) with multiple measurements on the same substrate (spot-to-spot) as well as comparing between different substrates (batch-to-batch), with values of 2.37% and 3.34%, respectively (n = 19).\textsuperscript{99} Su et al.\textsuperscript{256} report a RSD of 7.8% using AuNP clusters in a microfluidic channel (Figure 6A(v)). Microfluidic SERS, of likely use for in-the-field saliva analysis, can circumvent a well-known reproducibility problem in SERS, where it is difficult to control the distribution of analyte molecules across the SERS-active area. This is achieved by measuring the analyte in the aqueous phase.\textsuperscript{241} However, this does not mitigate any inherent variability in the plasmonic properties of the SERS substrate due to imperfections in the nanostructured morphologies. Reusable SERS-based microfluidic devices have appeared for potential in-field application (Figure 4A),\textsuperscript{47,146} however, typically, SERS substrates are not reusable, with analytes adsorbing to metal surfaces, laying the need for developing cost-effective disposable SERS substrates.\textsuperscript{85,308,336}

### 8.2. Raman versus SERS

Differences can exist between Raman and SERS spectra. This can include orientation effects relative to the incident radiation induced by the nanostructure topographies,\textsuperscript{296,337} or affinity molecules with specific moieties being more proximal to the high local electric fields\textsuperscript{11} or alterations to the Raman polarizability, which can increase the Raman cross-section by bringing the vibrational resonance closer to the laser excitation wavelength.\textsuperscript{87} Colceriu-Şimon et al.\textsuperscript{96} studying saliva and gold nanoparticles on a CaF\textsubscript{2} substrate, notice significant differences between saliva spectra for Raman and
SERS where the most prominent peak of thiocyanate at 2107 cm$^{-1}$ is only apparent in the SERS spectrum, which the authors attribute to a high affinity of the thiocyanate ion to metallic surfaces (Figure 6C). Contrariwise, no such effects appear to be prominent in a SERS study of overdose drugs in saliva, where a Raman database is used to identify the different compounds.\textsuperscript{[50]} Shende et al.\textsuperscript{[141]} record preferential adsorption in a SERS study of cocaine, caffeine, and phenobarbital, depending on the metal chosen as the plasmonic enhancer, while other drug studies make similar observations.\textsuperscript{[11,50,338]} More generally, the variable affinity of various compounds to the plasmonic metal, usually gold or silver to provide significant enhancements in the visible range, may be underappreciated in SERS.\textsuperscript{[48,339]}

8.3. SERS-saliva studies with functionalized surfaces

Durucan et al.\textsuperscript{[335]} purport that non-functionalized SERS is rare in the real world, citing lower sensitivity and quantification difficulties in multi-component samples where non-specific binding of interloping substances and competitive adsorption is problematic. For instance, an aptamer-functionalized gold SERS substrate has been used to detect a liver cancer biomarker in saliva at the nM range.\textsuperscript{[118]} Functionalizing SERS surfaces, however, adds system complexity, increases costs, and most of the saliva studies that utilize SERS use an unmodified SERS platform. Linker molecules create further challenges of preferential enhancement due to increased proximity to the SERS surface over the analyte molecules, such is the short-range nature of SERS (<10nm) and the precipitous drop in the electric field intensity away from the plasmonic surface.\textsuperscript{[340,341]}

8.4. SERS highlights from the Raman-saliva literature

The use of sol-gels to create a disposable device that can separate drug components in low volumes of saliva for SERS measurements has been extensively investigated by Farquharson and coworkers (Figure 6B(i)).\textsuperscript{[50,119,142,143,145,342]} The studies have demonstrated an ability to track chemotherapy drugs and illegal drugs to an LoD of 2 µg/mL for 5-fluorouracil and 50 ppb for cocaine, respectively.\textsuperscript{[50,143]} The sol-gels are formed by two precursor solutions, a silver amine complex and an alkoxide solution,\textsuperscript{[143]} mixed and spin-coated in a glass vial, followed by a reduction of the silver ions with sodium borohydride. The excess reagent is then flushed away with water. Remaining is the sol-gel, which is collected and packed into a capillary, providing NP stability. The sol-gel allows for the NPs to be trapped at a set size and aggregation. The small diameter and the fast measurement time of the capillary sol-gel device render it attractive for SERS measurements of saliva with merely 100 µL of the biofluid required, which can be collected and measured within 5 min.\textsuperscript{[142]}

Yuen et al. have shown Raman bands of saliva spectra to be preferentially enhanced via microwave heating (200 s, 600 W@2.45GHz), arising due to nanoscale changes to gold-coated polystyrene bead nanostructures (Figure 7A). Contrariwise, subsequent heating (600 s) appears to quench the intensity of certain spectral bands disproportionately.\textsuperscript{[193,194]} It is unclear whether this is purely a degradative photochemical\textsuperscript{[296]} or a plasmonic effect, where an alteration to the nanostructure topologies and roughness
affects the relative enhancement of the respective salivary Raman bands. Mohammadi et al. report SERS measurements of methamphetamine-doped artificial saliva focusing on the EF optimization of graphene-based materials on a dendritic silver SERS surface. Graphene oxide (GO) and gold nanoparticles (AuNPs) have been used to form a GO-AuNP nanocomposite SERS surface for application in a saliva study of gastric cancer. Here, rather than an electromagnetic or chemical SERS enhancing material, GO acts as a reductant in the AuNP synthesis, surfactant, as well as the supporting structure, aiding surface uniformity.

Zheng et al. show detection of silver (I) and mercury (II) ions in saliva via extrinsic SERS. In an intricate preparatory sequence, the authors fabricate a gold-covered titanium nanohole SERS surface pre-functionalized for single-strand DNA (ssDNA).
Concurrently, gold-silica core-shell nanostars are fabricated and functionalized for the complementary ssDNA strands. Raman dye reporter molecules are sandwiched between the silica shells (3 nm) and gold stars (~80nm). Subsequently, depending on the exact base pair sequence of the ssDNA used, the presence of Ag$^{+}$ or Hg$^{2+}$ can be confirmed due to these metal ions acting as intermediaries promoting the binding of the ssDNA strands. With the removal of the excess ssDNA and hence, the attached gold nanostars and the embedded Raman dye, the measured SERS signal indirectly indicates the concentration of metal ions present in the saliva sample. The gold nanohole substrate enhances the SERS signal by inducing strong plasmonic coupling with the nanostars.

Han et al.[346] have developed a SERS platform utilizing a novel DNA immunoassay. The platform is based on 23 nm diameter AuNPs functionalized with left and right DNA strands, with Malachite Green dye used as the Raman reporter. This assay, specifically designed to target the biomarker S100P associated with oral squamous cell carcinoma, demonstrates a LoD of 3 nM for detecting the isolated biomarker. Salmiliani et al.[47] present a AgNP SERS chip system employing dielectrophoretic aggregation to facilitate reusability. The authors show minimal nanoparticle fouling and chemical cross-contamination over three cycles (Figure 4A). Yang et al.[146] study a magnetically induced colloidal SERS assay for the detection of cotinine, a nicotine metabolite, and benzoylecgonine, a cocaine metabolite, in saliva (Figure 2B). In this reusable SERS system, consisting of Fe$_3$O$_4$ and AuNPs linked by IP$_6$ molecules, a RSD of 1.75% is reported using Rhodamine 6G.

The development of a microfluidic SERS sensor has been demonstrated by Sivashanmugan et al.[148] who use diatom bio-silica surface channels, which allow for trace detection of tetrahydrocannabinol. The porous channels are based on a diatom frustules substrate with AgNP grown in-situ on the surface. The substrate is placed in a microfluidic chip to chromatographically separate molecules from the biological complex with a SERS section for measurements. The diatom substrate has a porous surface enabling hot spots between the AgNPs in the pores, attributed to the achieved ultrahigh SERS sensitivity. The authors successfully identify tetrahydrocannabinol in saliva to a detection limit of 10$^{-9}$M in unprocessed saliva and 10$^{-12}$M in a water-diluted and centrifuged salivary sample.[196]

### 8.5. Numerical methods in SERS and saliva studies

Briefly, we also describe numerical studies within a Raman and saliva context, which are uncommon despite recognition that the development of SERS substrates is key to the progress of saliva in disease diagnosis.[64] In many cases, the SERS media used already have well-understood electromagnetic behavior i.e., spherical metal nanoparticles. In the broader SERS literature, numerical studies almost always model the local electric fields surrounding metal nanostructures, typically by the finite element method (FEM), finite difference time domain method (FDTD), or a discrete dipole approximation approach (DDA).[14,298] Following an initial substrate characterization study,[193] Yuen et al.[194] modeled the far-field extinction i.e., $-\ln($transmittance$)$, and SERS performance of microwave-irradiated polystyrene beads with DDA noting surface roughness and small gaps proved beneficial for larger enhancement (Figure 7A). These kinds
of effects are well-known in SERS and hark back to SERS’ electrochemically roughened electrode beginning.\cite{110,312,347}

While the behavior of evaporating droplets has been numerically modeled,\cite{285,348,349} and thus predictions on the spread of simple particulate matter can be made, this does not necessarily extend to more complex heterogeneous media, such as saliva and other biofluids. Further investigation in this area might prove useful. Analogously, as portable devices emerge, saliva studies would benefit from fluid flow models through microfluidic systems, which could inform on the dilution needed to achieve optimal channel passage i.e., viscosity, while mitigating detrimental effect on analytical sensitivity. For instance, Lee et al.,\cite{240} in a salivary microfluidic channel study into cotinine, a nicotine metabolite, simulate the fluid dynamics through the microfluidic geometry with FEM, noting that a drag force threshold exists that can impact antibody-antigen interaction.

Nanosystems combining photonic and plasmonic components have recently found sensing applications,\cite{350} notably whispering gallery mode sensors.\cite{351} In a Raman-saliva context, Hong et al.\cite{195} have presented an optical-plasmonic hybrid FDTD model consisting of an AuNP underlayer and a larger silica microsphere (Figure 7C). The authors measure methamphetamine-doped urine and saliva with the optoplasmic arrangement, presenting spectra down to concentrations of $1 \times 10^{-8}$M and $1 \times 10^{-9}$M, respectively. Zhang similarly explores an FDTD model but studies plasmonic coupling between a nanohole array and proximal nanostars.\cite{345} As an alternative application of numerical methods, Salemmilani et al.\cite{47} have investigated the electromigration of AgNPs in a dielectrophoretic SERS microchip with a FEM model, enabling the optimum location for SERS measurements to be determined and thus the identification of methamphetamine in saliva \textit{via} a chemometric model (Figure 7B).

Numerical methods can be computationally expensive, especially when high system fidelity in three dimensions is needed, and lately, reports have emerged, in nanostructure electromagnetic field and optics simulation,\cite{317,352–355} as well as in fluid dynamics,\cite{356,357} of using machine learning to improve efficiency while achieving suitably accurate results.

9. Data analysis

Datasets are increasingly large in modern science and the challenge is in the extraction of useful information.\cite{8,236} This is implied in the concept of “big data” with datasets set to be analyzed computationally in order to be analyzed meaningfully. Here, we summarize the standard spectra pre-processing methods, (univariate) statistical analysis, and the various multivariate techniques used in Raman-saliva studies. In a recent esophageal cancer study, Maitra et al.\cite{83} analyze, using Raman spectroscopy, not only saliva but also urine, serum, and plasma, and with the optimal quadratic discriminant analysis algorithm, saliva and urine are shown to provide the best classification accuracy, with saliva requiring the fewest selected variables from the spectra.

9.1. Spectra pre-processing

In the Raman-saliva literature, various pre-processing methods are used, including normalization, smoothing, cosmic ray and background removal (Figure 8). For baseline
subtraction, an asymmetric least square (AsLS) method\cite{83,103,155,223,239} or a high order polynomial fit\cite{97,99,220,227,244,358} are often applied. Normalization is performed using the area under the curve\cite{103,155,227,239} with several studies using peak normalization.\cite{217,219,220,223,227,244} Spectral smoothing has been carried out with the Savitsky-Golay filter,\cite{155,213,244} however, the window and the polynomial order are not always noted. Acquarelli et al.\cite{359} have shown that, depending on the specific sample interrogated (wines, beers, coffees, pharmaceutical tablets etc.), the technique used (Raman, IR), and algorithm employed, differing optimal pre-processing pipelines exist, which exclude some procedures and require those used to be applied in a certain order for the best classification accuracy i.e., there is no one-size-fits-all approach. Therefore, despite the similarities in pre-processing procedures in the current Raman-saliva literature, a saliva-specific investigation into optimal pre-processing pathways, with reference to the desired application, may be necessary.

A dimensionality reduction step may be further employed to reduce the computational load before a subsequent classification algorithm is applied. Often used in Raman-saliva studies is the unsupervised Principal Component Analysis (PCA) method,\cite{83,97,98,215,221,223,227,244,360} which separates unknown spectra by redefining the co-ordinate system to maximize the variation. Dies et al.,\cite{244} for instance, employ PCA to discriminate between different drugs in water and subsequently, spiked cocaine in saliva (Figure 4C). Hence, PCA may be used not only for dimensionality reduction but also classification outright. PCA permits the recognition of the most important peaks responsible for the variation between the inputted spectral data set, i.e., the PCA

![Figure 8. Procedures for Raman data processing and analysis. Red line is the most common data processing path among researchers in Raman-saliva studies*. A: data collection, B: baseline subtraction, C: normalization D: data cleaning in the form of dimensionality reduction/feature extraction E: classification F: results H: bypass of pre-processing steps. *Based on a subset of computationally focused Raman-saliva studies.](image-url)
loadings. The most important peaks in Raman-saliva studies can also be identified via genetic algorithms (GAs).\[83,103,239\]

### 9.2. Single variable statistical data analysis

Spectral features can be compared with statistical methods, primarily, hypothesis $p$-value testing for singular variables. The Student’s $t$-test\[153,215,217,219,225,227\] and (Wilcoxon-)Mann-Whitney $U$ test\[125,200,222\] have been used to study specific peaks or validate multivariate outputs. This analysis of singular variables allows for a strong depiction of the statistical relevance of a spectral feature adding to the data interpretation.

### 9.3. Multivariate data analysis

Traditionally, Raman spectra have been analyzed by selecting known peaks and monitoring changes in the intensity of these manually or with a classification algorithm. More recently, the whole spectrum has been investigated with multivariate techniques, with each datapoint representing a variable. Subtle differences between spectra can be identified and classification performed following in-feed of training data.\[332,361\] This comprehensive approach obtains information not just from peaks and troughs of the most important Raman bands but also uses the shape of these peaks alongside hitherto unappreciated variations across the entire spectrum. Therefore, for a complete analysis of complex biological processes in saliva, multivariate analysis is commonly recommended.\[8,61,86\]

In multivariate analysis, after initial PCA application, or feature extraction step otherwise, a supervised classification algorithm such as linear/quadratic discriminant analysis (L/QDA), or support vector machine (SVM) learning,\[2,83,99,155,204,244,362\] is employed. Partial least squares (PLS) analysis is a dimensionality reduction technique similar to PCA that additionally accounts for the correlation between the independent and dependent variables. It is also often paired with DA for classification purposes (PLS-DA) (Figure 3E).\[97,197\]

The use of algorithmic means is termed machine learning (ML), even often when unsupervised procedures are used. ML is frequently considered a subset of artificial intelligence and is used extensively in the Raman-saliva literature and ever more by the academic community at large. Increasingly common in classification is “deep learning” where feature extraction is a heavily layered algorithmic process used to further enhance the diagnostic performance.\[89,246,359,363\] Notably, deep learning includes artificial neural networks (ANNs), which mimic the neuronal networks in the brain.\[103,238,239\] Here, layers of interconnected nodes are activated to various degrees contingent on the data supplied. Inter-nodal connections, “synapses,” are assigned weights, which may be modified via a backpropagation algorithm designed to minimize error, a loss function, between predicted outputs and desired outputs.\[246\] Recently, ML has been highlighted as a valuable support tool to supplement medical diagnoses, reducing human error, across a wide range of medical fields,\[364\] even via patient-led smartphone applications.\[365–367\]
9.4. Sample size in Raman-saliva machine learning studies

In supervised ML the data acquired is split into test and training sets, and a sufficiently large number of training samples is usually required to have a meaningful model. Beleites et al.\cite{368} have recommended 75–100 samples per class for optimal training, based on computational experimentation. Unfortunately, acquisition of such large sample sizes is often problematic in healthcare studies, especially when subclasses of disease exist, for example, varying cancer stages.\cite{160} The importance of having a large enough sample size is to reduce overfitting, which corresponds to the model’s inability to generalize, i.e., the model develops bias toward the training samples, becoming unable to interpret samples from outside that of the training sample space. Test-training separations of 80%/20%\cite{83,125,132,239,369} are apparent in Raman-saliva studies, however, Raman studies on saliva have also been performed with 90%/10%\cite{204} and 60%/40% splits.\cite{362}

Within chemometric-focused Raman-saliva studies, there is a range of sample sizes used varying from 20 samples (10 healthy/10 unhealthy)\cite{96,103,215} to 128 (64 healthy/64 unhealthy) samples in total.\cite{132,238,360,369} Healthy/unhealthy patient splits usually are approximately equal. Radzol and coworkers gather data from a Raman spectral data-bank, which is easier to obtain than clinical samples from patients.\cite{132,360,369} As an important distinction, it is noted that large spectral datasets do not necessarily equate to a more accurate model diagnostically when obtained from a small sample cohort, i.e., few independent samples,\cite{368} although this may nevertheless assist in accounting for the inherent variation between the native salivary constituents. Recently, Guo et al.\cite{370} have proposed modified PCA and PLS algorithms accounting for (i.e., subtracting/mitigating) intra-class variance to improve feature extraction and classification accuracy.

The accuracy of the models is evaluated via assessment of the diagnostic sensitivity (true positive rate) and the diagnostic specificity (true negative rate), often in conjunction with receiver operating characteristic curves (Figure 2D(ii)(iii)). In conjunction with ensuring the reproducibility and accuracy of the acquired data sets, validation, i.e., testing with an externally acquired dataset, is further necessary and is generally unconsidered in Raman-saliva studies targeting specific diseases.\cite{236} In a Raman-saliva study into Dengue fever, Othman et al.\cite{360} have investigated the effect of principal component retention criteria (eigenvalue one criterion, cumulative percentage variance, and scree test) alongside varying numbers of ANN layers. Finally, we note that many other chemometric techniques are in use within the spectroscopy field\cite{371,372} and there is a need to compare the different algorithms on the same dataset.\cite{236}

10. Outlook

Several popular Raman approaches would appear to be overlooked in the Raman-saliva literature. First, there are no reports using tip-enhanced Raman spectroscopy (TERS), which uses scanning probe microscopy setups in combination with plasmonic materials to provide high sensitivity and resolution.\cite{82,373–375} This approach could be helpful to characterize specific salivary components. Second, no uses of hydrophobic surfaces are observed\cite{320,376–380} or surfaces otherwise deliberately functionalized to alter the wetting properties of drop-cast saliva and the subsequent deposition pattern of salivary components by, for instance, mitigating contact line pinning in a drying microdroplet.\cite{270} Su
et al.\textsuperscript{256} hydrophilize (through hydroxylation) microfluidic glass channels \textit{via} Piranha solution treatment followed by a 12 hour NaOH immersion.

A single study of shell-isolated nanoparticle enhanced Raman spectroscopy (SHINERS) is presented by Al-Ogaidi et al.,\textsuperscript{226} detecting glucose in saliva with SiO$_2$-covered AuNPs. SHINERS are core-shell nanoparticles for SERS consisting of a plasmonically active core material surrounded by a thin, inert shell, often silica, ensuring that the inner material does not participate in any adsorption with the analyte.\textsuperscript{381} Given the potential interferents in saliva, or indeed any bio-matrix, their exclusion from the literature at present might seem surprising. Perhaps most surprisingly, however, resonance Raman effects are seemingly seldom employed.\textsuperscript{105}

11. Conclusions

Saliva is a promising biofluid for use in bedside, roadside, pitch-side, or other point-of-need settings. Combined with Raman spectroscopy, faster, more accurate, and multiplexed determinations could be made, whether the analytes be physiological or pharmaceutical in origin. In this review we have reported the common methods employed in the collection, pretreatment, and storage of saliva within Raman studies, alongside subsequent aspects of the measurement procedures, and Raman parameters used. Given the need to detect low concentrations or subtle changes in compounds of interest, we have further surveyed SERS methods employed and summarized data analysis techniques.

A variety of methods are evident across the various experimental facets, and hence we conclude there is little adoption of standard procedures. In some, such as saliva storage, good guidelines exist elsewhere, and these could be fine-tuned to the needs of the spectroscopist or with specific applications/analyte molecules in mind. For other aspects, such as the use of drop-casting and the effect of measurement on inhomogeneous surface distributions, there will need to be more careful consideration within a saliva context. SERS is commonplace within the Raman-saliva literature, although a greater focus on substrate development, including numerical models, may be required. A wide range of chemometrics is employed, and more comparisons of these techniques side-by-side in the same saliva study would be useful. Many of these concerns may, in fact, be indicative of the emergent nature of the field, with majority of studies being preliminary. With the standardization of collection, storage, and Raman measurement protocols for saliva, coupled with technological development into portable, even handheld, devices, advances in data science, as well progress in SERS substrate development, Raman-saliva can become a biofluid-spectroscopy combination for real-world applications.

Notes

1. In 2020 a fourth pair of salivary glands, the “tubarial glands,” found in the \textit{torus tubarius} section of the nasopharynx, were identified \textsuperscript{382}. The designation of these structures as a salivary gland is being disputed.

Acknowledgments

We acknowledge funding from the Wellcome Trust (174ISSFPP), the Royal Academy of Engineering (RF1415\14\28), the Defence Science and Technology Laboratories (DSTLX-
Disclosure statement

The authors declare no conflict of interest. The authors do not endorse nor have any relationship with any of the companies mentioned herein.

Author contributions

MH, LK and PGO conceptualized the review. MH drafted the review and performed the meta-analyses. MH, LK, PDCG, EB & HOMC collated data from the literature. LK collated data for Tables 1 and 2. PDCG contributed to the section on data analysis. LK contributed to sections on data analysis and SERS. PDCG analyzed chemometric trends. EB provided input on microfluidic devices and HOMC on dental/orthodontic studies. MC contributed to obtaining rights and permissions. MH and PGO were in charge of the overall direction and planning, supervised the review, guided, revised and overall edited the manuscript. All authors have given approval to the final version of the manuscript. All authors discussed the review and commented on the manuscript.

ORCID

Mike Hardy https://orcid.org/0000-0002-0242-8727
Liam Kelleher https://orcid.org/0000-0001-6392-8875
Paulo de Carvalho Gomes https://orcid.org/0000-0003-2986-1157
Hin On Martin Chu https://orcid.org/0000-0002-1968-919X
Pola Goldberg Oppenheimer https://orcid.org/0000-0002-1014-4724

Data availability statement

All data used in meta-analyses in Figure 3 is available in Table S1 in Supporting Information.

References


120. Frosch, T.; Knebl, A.; Frosch, T. Recent Advances in Nano-Photonic Techniques for Pharmaceutical Drug Monitoring with Emphasis on Raman Spectroscopy. Nanophotonics 2019, 9, 19–37. doi:10.1515/nanoph-2019-0401


331. Stensberg, M. C.; Wei, Q.; McLamore, E. S.; Porterfield, D. M.; Wei, A.; Sepúlveda, M. S. Toxicological Studies on Silver Nanoparticles: Challenges and Opportunities in Assessment, Monitoring and Imaging. *Nanomedicine (Lond.)* 2011, 6, 879–898. doi:10.2217/nmm.11.78


339. Ye, Z.; Li, C.; Xu, Y.; Bell, S. E. J. Exploiting the Chemical Differences between Ag and Au Colloids Allows Dramatically Improved SERS Detection of “Non-Adsorbing” molecules. *Analyst* 2019, 144, 448–453. doi:10.1039/c8an01927d


361. Wold, S. Chemometrics; What Do We Mean with It, and What Do We Want from It? Chemom. Intell. Lab. Syst. 1995, 30, 109–115. doi:10.1016/0169-7439(95)00042-9


