Using Raman spectroscopy to characterise biological materials

Holly J. Butler[1,2], Lorna Ashton[3], Benjamin Bird[4], Gianfelice Cinque[5], Kelly Curtis[6], Jennifer Dorney[6], Karen Esmonde-White[7], Nigel J. Fullwood[8], Benjamin Gardner[6], Pierre L. Martin-Hirsch[1], Michael J. Walsh[9], Martin R. McAinsh[1*], Nicholas Stone[6*], Francis L. Martin[1*]

[1]*Lancaster Environment Centre, Lancaster University, Lancaster, UK;* [2]*Centre for Global Eco-Innovation, Lancaster Environment Centre, Lancaster University, Lancaster, UK;* [3]*Department of Chemistry, Lancaster University, Lancaster, UK;* [4]*Daylight Solutions, San Diego, CA, USA;* [5]*Diamond Light Source, Harwell Science and Innovation Campus, Chilton, Oxfordshire, UK;* [6]*Biomedical Physics, Physics and Astronomy, University of Exeter, Exeter, UK;* [7]*Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA;* [8]*Biomedical and Life Sciences, School of Health and Medicine, Lancaster University, Lancaster, UK;* [9]*Department of Pathology and, Department of Bioengineering, University of Illinois at Chicago, Chicago, IL, USA.*

* Corresponding authors: Prof. Francis L Martin PhD, Centre for Biophotonics, Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK; *E-mail*: f.martin@lancaster.ac.uk; *Tel.*: +44 (0)1524 510206; Dr Martin R. McAinsh m.mcainsh@lancaster.ac.uk; Prof Nicholas Stone n.stone@exeter.ac.uk

Co-author e-mail addresses:

Lorna Ashton - l.ashton@lancaster.ac.uk

Holly J. Butler - h.butler@lancaster.ac.uk

Benjamin Bird - bbird@daylightsolutions.com

Gianfelice Cinque - gianfelice.cinque@diamond.ac.uk

Kelly Curtis - kmc217@exeter.ac.uk

Jennifer Dorney - J.Dorney@exeter.ac.uk

Karen Esmonde-White - kareshni@umich.edu

Nigel J. Fullwood - n.fullwood@lancaster.ac.uk

Benjamin Gardner - B.Gardner@exeter.ac.uk

Pierre L. Martin-Hirsch - martin.hirsch@mac.com

Michael J. Walsh - walshm@uic.edu

ABSTRACT

Raman spectroscopy is an optical technique that measures the chemical composition and molecular structure of a sample. Utilisation of the "molecular fingerprint" of Raman spectra has proven an effective analytical approach in geology, semiconductor, materials and polymer science fields. The application of Raman spectroscopy and microscopy within biology is rapidly increasing and it has proven to be a valuable analytical tool for various applications. The technique does not typically suffer from interference from water molecules, and therefore biological materials are ideally suited to this spectroscopic approach even with interference from tissue auto-fluorescence. Analysis does not conventionally require extensive sample preparation and allows for the label-free extraction of biochemical and structural information. We describe a protocol for acquiring Raman spectra, maps and images from a range of biological samples, and explore a robust approach for sample preparation, instrumentation, acquisition parameters and data processing. In this protocol we aim to standardise and bring together multiple experimental approaches from key leaders in the field. Using this standardised approach, we expect that a typical Raman experiment can be performed by a non-specialist user to generate high quality data for biological materials analysis.

INTRODUCTION

Raman spectroscopy has been shown to be a powerful analytical technique in the study of biological materials, allowing rapid, non-invasive and high spatial resolution acquisition of biochemical and structural information through the generation of point spectra or spectral images. Although traditionally used for analytical chemistry applications, there has been a significant rise in the use of this technique within biological studies, particularly in the field of biomedicine[1-4].

Raman spectroscopy

Raman spectroscopy utilises monochromatic light, often in the near infra-red (NIR), visible or ultraviolet (UV) range, to exploit the phenomena of inelastic scattering, or Raman effect, that describes the excitation of photons to virtual energy states and the resultant loss (Stokes) or gain (anti-Stokes) of energy that occurs due to the interaction of light with vibrational modes associated to chemical bonds within the sample. This shift in energy is indicative of discreet vibrational modes of polarisable molecules, and thus a qualitative measurement of biochemical composition can be obtained. Raman spectra can infer quantitative information provided that the instrument response function is adequately corrected. Typically, the significant regions of the Raman spectrum that are observed within biological specimens fall within 400 - 2000 cm^{-1} wavenumbers, associated with bond vibrations of proteins (1500 - 1700 cm^{-1}), carbohydrates (470 - 1200 cm^{-1}), phosphate groups of DNA (980, 1080 and 1240 cm^{-1}) and additional cellular biomolecules[1,5]. Higher frequency bond vibrations associated with CH, NH and OH stretching in lipids and proteins can also be observed at higher wavenumbers (2700 - 3500 cm^{-1})[5]. Consequently, a distinctive biological 'fingerprint' is derived from the biological sample under investigation and can contribute to our understanding of the specimen. **Figure 1** illustrates the principle of elastic (Rayleigh) and inelastic scattering when analysing biological specimens.

Raman spectroscopy has a high molecular specificity, making it an excellent technique for materials analysis. However, Raman scattering is a rare phenomenon with an exceptionally low probability of occurrence (~1 in 10^8) in comparison with its counterpart Rayleigh scattering, and thus is typically described as an insensitive technique. Its increased applicability is largely attributed to technological advancement of highly efficient laser sources, low-noise detectors, effective Rayleigh filters, and high-throughput optics[6]. The development of microspectrometers, combining the power of optical magnification and direct visualisation of the sample with the derivation of corresponding spectral fingerprints, have also contributed to further exploitation in biological fields. Due to this, it is possible to not only derive single point spectra, but also to produce highly informative Raman images of the sampling area with improved interpretability[7]. The ability to acquire high quality spectra at sub-cellular resolution coupled with the capability to obtain valuable information non-invasively, label-free and without interference from water, makes Raman spectroscopy an ideal approach for *in vivo* biological investigations.

There are many factors that influence spectral resolution which are outlined in detail in this protocol. Whereas, the spatial resolution of optical Raman microspectroscopy is governed mainly by the diffraction limit of light, and therefore is dependent upon the laser wavelength (l) in use as well as to the numerical aperture (NA) of the objective. Theoretically, for a

confocal microscope the diffraction limit for visible (l - 488 nm) to near IR (~1033 nm) and from long-working distance to water/oil immersion objectives with NA from 0.5 - 1.2 would be in the range of 0.2 to 1 micrometre. In practise, it is rarely possible to achieve the diffraction limit both due to imperfect optics and beam scattering at the interface of the sample. However, the use of alternative sampling modes such as surface-enhanced Raman scattering (SERS) and tip-enhanced Raman spectroscopy (TERS) have been shown to provide spatial resolution below the diffraction limit as targeted single molecule detection is possible[8,9]. This is due to the interaction of biomolecules with roughened surfaces, such as metallic nanoparticles, that effectively enhance the electric field by a factor of up to 10^{14}. This is associated in part to the excitation of localised surface plasmons, known as the electromagnetic theory, as well as the charge transfer between the two materials.

A typical Raman study can rapidly accumulate a large, information-rich spectral dataset. However, as this dataset expands, the extraction of biological information becomes increasingly challenging. For this reason, multivariate analysis approaches are often employed in order to effectively extricate the underlying chemical and structural information[10]. Spectral datasets often present a significant computational burden and thus analysis usually includes a data reduction step, particularly as there are many covariant features in the spectra. This information can then be fed into unsupervised (clustering) or supervised classification to differentiate individual spectra, which can then infer biological information[11]. Raman imaging (global illumination) and mapping (step-wise) have particularly benefitted from these classification capabilities, as spectral discrimination is indicative of the underlying biological architecture, which has proved valuable in the field of cancer diagnosis, where this approach has assisted conventional histopathological techniques[12].

Applications

Raman spectroscopy presents a method of sample examination with a high degree of flexibility enabling data to be recorded from a diverse array of sample types including fixed, fresh or live tissues and cells. The application of Raman spectroscopy to characterise biological materials is a rapidly expanding field and has been used in the field of pharmacology[13-15], microbiology[16-20], toxicology[21,22], plant science[23-25] and human biology. Considerable advances have been made, particularly in regards to cancer diagnosis and prognosis[26], with clinical implementation firmly on the horizon[27], spearheaded by recently formed networks such as 'Raman4Clinics' (www.raman4clinics.eu/) and Clinical Infrared and Raman Spectroscopy for Medical Diagnosis (CLIRSPEC - www.clirspec.org/). Raman spectroscopy and its derivatives coupled with multivariate analysis can classify neoplasia in brain[28-31], breast[32], bladder[33], colorectal[34,35], larynx[36], lung[37], lymph nodes[38,39], oesophageal[40-42], prostate[43-46], uterine and cervical[47-50] tissues using both point spectra and imaging approaches. These examples have employed a range of sample formats, including fixed cells and tissues, *in vivo* measurements with advances in fibre optic attachments and spatially offset Raman spectroscopy (SORS), as well as non-invasive biofluids measurements[51]. In addition to cancer studies, the technique has been employed to shed light upon infectious diseases such as malaria[52-55]. Due to the relatively fast sampling time and mapping ability of Raman microspectroscopy in aqueous environments, the 2D and 3D analysis of *in vitro* models is possible[56-59], which has been shown to be particularly useful in toxicology and therapeutic studies[60,61].

The application of SERS has been shown to vastly improve the spectral intensity obtained from typically weak Raman scatterers, such as dilute biofluids[62]. SERS has been employed across a wide range of biomedical studies, including DNA and drug detection[63-66]. This approach has also been shown to overcome autofluorescence, which has often limited research capabilities in samples with intrinsic fluorescence, like chlorophyll in plants, as this strong signal can often engulf the relatively rare Raman event[24]. Reduced Raman performance in plant research has significantly restricted its application in this area, with previous research focused on quantification and identification of plant constituents[67-72], as well as imaging of non-fluorescent tissues[25,73-75]. As fluorescence has a defined wavelength profile, it is also possible to overcome fluorescence by employing alternative radiation sources using lasers at wavelengths outside typical biological matrix absorbance e.g. in the NIR (1064 nm) region[76]. Water has also been found to be a significant fluorescence quencher, which has benefitted *in vivo* studies, recently demonstrated in fundamental plant monitoring[77,78].

Limitations

The molecular specificity of Raman spectroscopy is powerful for the study of biological materials. However, there are a number of disadvantages associated with the technique. Since Raman scattering is a relatively low probability event; low sensitivity can be an issue, which is exacerbated by interference from fluorescence[13]. Due to the intense laser powers typically used, local thermal decomposition of the sample may be encountered, especially when using UV or visible wavelengths. Sample burning (or photoablation) may be observed visibly by dark areas on the sample and spectroscopically by the presence of amorphous carbon bands (C=C ~1500 cm^{-1}) in the spectra, or saturation of the detector. The protocol described here will provide troubleshooting advice to avoid these particular issues. Fourier transform infrared (FTIR) spectroscopy is an alternative vibrational spectroscopy technique that is also suitable for analysis of biological materials (for a recent protocol, see Baker et al[79]). Although based on fundamentally different physical excitation processes, FTIR retains both molecular specificity and sensitivity, and has thus been employed to derive spectral data and elucidate biological information, complementary to Raman spectroscopy[80-82].

The aim of this article is to describe a specific protocol for Raman spectroscopy that can be applied to a variety of biological samples independent of the specific manufacturer instrumentation and software. The protocol will address (i) sample preparation, (ii) spectral acquisition, and (iii) data analysis of spectral datasets with anticipated results derived from a range of biological studies (**Fig. 2**). It is important to note that this technique can also be applied to non-biological materials, providing insights into the steps involved in Raman studies. In this protocol we approach principal aspects of planning and implementing Raman investigations that can be applied to a variety of biological samples. We direct the reader to additional protocols that approach specific biological applications[75,83-86]. It is our aim that by unifying protocols from leading researchers in the field, spectroscopists and biologists can build new interdisciplinary studies into biological samples incorporating Raman spectroscopy into the suite of molecular biology tools.

Experimental design: instrument options

When conducting a biological study using Raman spectroscopy it is important to first consider the investigative aims (exploratory versus diagnostic) and the desired analytical goals (imaging, classification for diagnosis, pattern finding, and biomarker identification)[79,87]. Furthermore, the constraints of the sample under investigation must be addressed (e.g., format, impurities) and the consequent suitability of different Raman substrates (e.g., quartz, metal-coated glass, calcium fluoride). **Figure 3** presents an overview of a typical Raman system and identifies key instrumental components. There are numerous Raman spectroscopic variations available, as described in **Table 1**. This protocol will focus primarily upon spontaneous Raman microspectroscopy.

Excitation source. It is possible to use a variety of excitation sources to irradiate samples, dependent on the suitability of the specimen to increased photon energies and the sensitivity and resolution required in the study. Broadly speaking there are three main factors to consider when choosing a laser system: (i) source, (ii) wavelength, and (iii) spot size. There are several laser source options available within Raman systems, of which diode lasers are increasingly commonplace, particularly within biological systems. Diode lasers are based upon solid or semi-conductor technology and consequently provide greater energy efficiency than their popular noble gas-based counterparts[88,89]. Although gas-based lasers have been frequently used in laser-based technologies, they are restricted by a finite lifetime. A helium neon (HeNe) laser has a relatively low power output in comparison to diode lasers, and more powerful alternatives, such as the argon-ion laser, discharge large amounts of heat that requires a cooling system to counteract, which is energy inefficient and also reduces portability[88]. The introduction of diode-based lasers has contributed to the development of portable Raman systems due to their durability and compactness[90,91]. When choosing a laser source, it is also important to consider if the experiment requires the use of pulsed (e.g., neodymium-doped yttrium aluminium garnet [Nd:YAG]) rather than continuous wave lasers (e.g., Krypton ion laser). Line width of the laser is also an important consideration as it has a direct influence upon the spectral resolution, regardless on spectrometer configuration. Generally, gas-based and solid-state lasers have narrow bandwidths, whereas some multi-mode diode lasers can have wide bandwidth with broad spontaneous emission.

The wavelength of the laser is critical to the experimental design (**Fig. 4**). Wavelengths available are laser specific, yet generally fall between the mid UV (>200 nm) and the NIR (<1.1 μm). As mentioned previously, it is possible to suppress fluorescence using NIR wavelengths, due to the use of relatively low energy photons. Similarly, UV wavelengths can be employed as the fluorescence is red-shifted beyond the fingerprint region of the Raman spectrum, and therefore is comparatively free from fluorescence, despite much higher energy wavelengths[76]. A key constraint to consider is the interaction between the laser wavelength and the sample, due to the effects of phototoxicity. High energy or prolonged exposure to light can have a destructive effect on samples, particularly *in vivo* studies and interrogation of sensitive samples such as single cells. It is advised to reduce exposure times wherever possible and to conduct adequate controls to determine toxic effects of the laser wavelength in use.

In addition, it is important to consider the effect that laser wavelength has on other key experimental parameters such as resolution and sensitivity. The spatial resolution is

dependent on wavelength of the laser, due to the determination of the laser spot size. Sensitivity is also dependent upon the excitation wavelength, as the Raman scattering intensity is proportional to $1/\lambda^4$ resulting at significantly reduced sensitivity at lower laser photon energies, such as the 1064 nm laser wavelength.

NIR lasers, most commonly 785 nm and 830nm, have been extensively applied in biological studies, particularly fixed and live cells, as the lasers have relatively low photon energy and generally do not cause substantial photodamage[92-95]. Tissue Raman spectroscopy is also typically performed in the NIR because it is within the diagnostic window of low melanin and water absorption[96]. Excitation in the UV region has been particularly useful in resonance Raman studies of biological materials, due to the similarity between the laser wavelength frequency and the electronic transition of the molecule under investigation[97,98]. The total laser power onto the sample is central for acquiring high quality results from biological samples given that they are generally low scattering materials and suffer from radiation damaging. Practically, it is better to have >30 mW laser power at the source for the shorter visible wavelengths and >100 mW for the longer ones, then use a series of attenuators to optimize the illumination power delivered at the specimen. It is important to consider this last point, as the laser intensity is dependent upon laser spot size (sampling area) and magnification and can have a significant impact on laser exposure on the sample.

It is possible to alter the profile of the laser spot in some Raman systems, which can also be beneficial depending on the experimental parameters and overall investigative aims. By focusing the laser in a line, a larger surface area is illuminated, maximising the Raman scattered light, which in turn derives spectral information from a larger surface area as well as preventing photodamage[99]. Line mapping by raster scanning the line focused laser across the sample, can obtain Raman maps with improved acquisition times, without compromising image quality[83,100-102]. The intensity across the laser line is variable and therefore this approach leads to spectra with variable signal to noise across the laser line. An alternative is to globally illuminate the sample area, resulting in high quality spectral imaging with enhanced power distribution across the sample[103,104]. Specialised filters (e.g., liquid-crystal tunable, dielectric, acousto-optic filters) are used to actively select specific wavelengths, which along with a 2D detector allows the production of true Raman images from relatively flat samples[101]. Although global illumination allows for high quality spectral imaging, this often comes at a cost of spectral quality due to the reduced laser intensities across the sampling area.

Detector. In order to detect the weak intensity of scattering, the detector included in the Raman system needs to be extremely sensitive. Greater levels of sensitivity may be necessary in some studies where scattering is especially weak, or where noise is particularly high and thus the challenge of detection is increased. Charge coupled devices (CCDs) are commonly integrated in Raman systems because they exhibit high quantum efficiencies and low signal-to-noise (SNR), compared to early alternative detectors such as photomultiplier tubes (PMT) and photodiode array (PDA)[105]. CCDs are multichannel arrays made up of made up of thousands of pixels, each of which can collect charge from scattered photons[106]. This charge is then passed through adjacent pixels, until they are readout individually, relating directly to the scattering intensity. There are a range of CCD choices available, including intensified CCDs (ICCDs) and electron multiplying CCDs (EMCCDs). At extremely low scattering intensities, EMCCDs can provide superior sensitivity than conventional CCDs due to creation

of further electrons and therefore providing a signal relatively higher than the readout noise. Although, if spectra are shot or Poisson noise limited, the use of an EMCCD will not improve SNR[107]. Dark, or thermal, noise can be dramatically improved by detector cooling, often using liquid nitrogen cryogenic or thermoelectric Peltier cooling, with deep cooling toward -80 - 100°C providing up to one order of magnitude improved noise reduction[108]. The quantum efficiency of the silicon based CCD detectors is wavelength dependent and drops off rapidly in the NIR region. In back-illuminated CCDs this can result in fringing effects as the detector effectively behaves as an étalon. Longer wavelengths of light that are not effectively absorbed by the photosensitive region of the detector are instead reflected throughout the detector, with detrimental effects upon spectral quality. This can be somewhat overcome by using deep depletion CCDs. For NIR studies beyond 950 nm, where the photon energy is less than the silicon bandgap, NIR detectors such as an indium gallium arsenide (InGaAs) detector are advantageous to overcome thermally generated noise in the higher wavelength region[109].

Additional hardware. Rayleigh scattering is more intense than Raman scattering and can easily overpower the more informative signal so it must be optically filtered. Dispersive Raman spectrometers use specialised Rayleigh filters or a multistage monochromator, whereas non-dispersive spectrometers often employ Fourier transformation (FT) based on a Michaelson interferometer[109]. There are a range of manufacturer-specific Rayleigh filters available[110], however holographic notch and dielectric edge filters are most commonly used. Edge filters transmit light wavelengths above that of the laser in use, whereas notch filters will effectively filter only the laser wavelength, allowing both Stokes and anti-Stokes measurements[111]. Metal oxide edge filters have been shown to have much longer lifetimes than notch filters. Rayleigh filters must be selected to be specific to the laser wavelength. In contrast, a multistage monochromator can be used with variable laser wavelengths although there is a major throughput disadvantage to using multistage monochromators[6]. Single monochromators are comprised of a diffraction grating, used to disperse the Raman scattered light and are universally used in conjunction with Rayleigh filters. Gratings differ due to the number of grooves (g/mm) or lines (l/mm) on the surface, which can be anywhere in the range of 150 - 4000 per mm, corresponding to the diffraction, or angular dispersion capabilities of the grating. Higher groove frequency can improve the spectral resolution at the cost of reduced spectral intensity and range. The sampling aperture of the system determines how much light, and therefore Raman scatter, is passed through the spectrometer. Slits and pinholes range from 10 μm to 100 μm, with a larger aperture allowing more light through the system and thus increasing sensitivity, but at the expense of spectral and depth resolution. For thin samples liked fixed cells the use of pinhole and optimal in-depth focus of the laser illumination is crucial to maximize the Raman signal from the sample volume.

The choice of microscope objective is crucial in Raman microspectroscopy. The objective's throughput depends on its magnification and solid angle of light collection i.e. NA. A high magnification/high NA objective provides higher axial spatial resolution but has a shorter working distance and may be better suited for thin samples. Low magnification/moderate NA objectives provide lower axial spatial resolution but have a longer working distance and may be better suited for bulky specimens to avoid the specimen touching the objective. Low magnification may be more appropriate for samples that do not require high spatial resolution, such as liquids and homogenous samples. Conversely, high magnification is beneficial in studies that require high spatial resolution so that specific biological architecture

can be examined. Importantly, microscope objectives can be specially coated for optimal performance in a specific wavelength region and those coatings can produce significant background signals if another wavelength is used. We encourage careful consideration of the microscope objective and additional preliminary tests with the sample under investigation.

Experimental design: sample preparation

One of the major advantages of Raman microspectroscopy in biological studies is the ability to derive label-free and non-destructive spectral information, with minimal sample preparation. However, it is important to appreciate sample constraints as well as substrate options that can have a significant impact on experimental procedures. As the technique relies upon maintaining optical focus, sample stabilisation and relative flatness are important sample preparation considerations.

Sample format. Fixation can be an important step in sample preparation and is often used to isolate a sample at a distinct time point or for sample archiving. Formalin-fixed, paraffin-embedded (FFPE) tissues have been historically archived in pathological settings and specimens have been widely analysed using microspectroscopy. Formalin modifies the proteins by cross-linking and paraffin has strong signals in the fingerprint and higher wavenumber region of the spectrum. These signals can be found at 892, 1065, 1135, 1174, 1298, 1421, 1443 and 1464 cm^{-1} and thus have significant overlap with the underlying sample biology[112]. The contributions of paraffin in the Raman spectrum are removed by either de-waxing of the sample, which has been shown to have dramatic effect on lipid content, or digital de-waxing that can remove the strong paraffin peaks from the spectrum[113]. Analysis of fresh or snap-frozen tissues may overcome these substantial drawbacks although sample acquisition becomes more difficult, sectioning is encumbered, and sample degradation must be controlled. Lipid alterations are also evident in fixed cells analysed using Raman spectroscopy[114]. Live cell analysis is therefore a rapidly expanding field that allows the user to interrogate cells *in situ* in aqueous environments and 3D cell cultures whilst maintaining key growth parameters including temperature and gas availability[56]. In these studies cells should be sufficiently adhered to a growth substrate prior to analysis. Also, the background signal from the cell media should also be considered prior to analysis as it may provide unwanted background interference.

Liquid samples can be easily examined using an immersion objective or a microfluidic device to overcome spectral artefacts due to the presence of bubbles and surface tension. This has permitted the analysis of biofluids, such as blood plasma, sputum, saliva and urine, in diagnostic studies[51]. Another approach for biofluids analysis is a combined drop coating deposition Raman spectroscopy (DCDRS), also known as drop deposition/ Raman spectroscopy. Small biofluid volumes are deposited onto a flat substrate and allowed to dry based on sessile drop formation principles[115]. Sessile drop formation has many benefits for Raman spectroscopy with significant pre-concentration of proteins, coarse separation of impurities and reproducible prediction of protein solution concentrations[116,117]. The resultant "coffee ring" dried drop can be examined by Raman microspectroscopy, with the knowledge that a variable coffee ring thickness affects spectral intensities and distribution of macromolecules within the ring deposit[118-120].

Raman substrates. Substrate choice is a critical factor in experimental design that is dependent upon the experimental outputs and sample characteristics. The matrix upon which

9

a sample is supported contributes to stability and therefore directly affects spectral quality by maintaining focus. The most significant properties of the substrate to consider are the spectral background signals, the substrate cost, availability and composition. Glass slides typically used in optical microscopy are exceptionally cost-effective but have a strong background fluorescence that limits analysis at excitation wavelengths other than 532 nm[121]. Another approach is to use metal-coated glass slides, such as aluminium or gold-coated glass, which effectively eliminates the glass signal. Gold-coated slides with roughened surfaces have also been shown to be good SERS substrates[122]. Calcium fluoride (CaF_2) and quartz slides are used as Raman substrates due to minimal background interference. Care should be taken in choosing a substrate because contaminants can produce unwanted backgrounds. Barium fluoride (BaF_2) slides have also been used as a Raman (and IR) substrate but their partial solubility in water make them unsuitable for *in vivo* and aqueous studies[123]. Quartz and fused silica are also popular choices for Raman studies. While these speciality substrates are more expensive than glass, they are often reusable in a laboratory setting. In a clinical setting, these costs may be prohibitive, particularly for large scale screening programs, and we support development of low-cost, single-use sterilised Raman substrates.

Experimental design: spectral acquisition

The desired spectral output must be considered when approaching spectral acquisition, as experimental parameters may be altered between simple point spectra and image construction. In general, superior spectral quality can be obtained using a point spectrum approach, as datasets can benefit from longer acquisition times. In comparison, hyperspectral data cubes generated from mapping and imaging techniques can provide user friendly data interpretation, although negotiation between extensive acquisition times, spectral and image quality is required[101]. For example, a map spanning a biological feature 10 x 10 μm in size, would be compromised of 100 spectra when using a step size of 1 μm (with over-fitting). This map could be acquired relatively quickly when using acquisition times less than 10 seconds, although some samples may require increased laser exposures due to small volumes or weak scattering tendencies such as single cell studies. It is important to note however, that huge numbers of poor quality spectra can still provide exquisite detail using multivariate approaches.

Resolution. The desired spectral and spatial resolution of the experiment, and hence the required sensitivity, must be decided prior to spectral acquisition. As mentioned with regards to instrumentation options, spatial resolution is ultimately diffraction limited and therefore linearly dependent on the laser wavelength and inversely with the objective NA. Accordingly, high spatial resolutions can be achieved with lasers at shorter wavelengths and high magnification optics.

Spectral resolution is in fact greater at higher excitation wavelengths, providing the Raman optical configuration remains constant. Gratings with higher groove frequency can improve spectral resolution at a cost to sensitivity and spectral range. An increased number of smaller-sized pixels within the detector and reduced between pixel cross talk, can also contribute to increased spectral resolution[124]. However, this is often impractical to alter within a single Raman system. Spectrometer focal length (SFL) refers to the distance between the diffraction grating to the detector typically between 200 - 800 mm, with greater distances generally providing improved spectral resolution[125]. A larger SFL requires a larger entrance slit, in

order to allow the maximum passage of light, which can also influence spectral resolution[126]. Although not exhaustive, these factors determine the spectral output and must be considered prior to analysis.

It is evident that when choosing a laser excitation wavelength there is a potential trade-off between desired spectral and spatial resolution. High spectral resolution (up to 0.5 cm^{-1}) may be required in studies where specific molecular information is required and thus adjacent Raman bands need to be differentiated, such as in pharmaceutical studies monitoring drug uptake at the cellular level. Increased spatial resolution (up to 1 μm) is required when specific localised information is required such as when interrogating tissue features to extract biological information. This is particularly valuable in mapping and imaging approaches as specific architecture can be imaged, such as when identifying cancer progression in tissue samples[42,45].

Calibration. Commercially available instruments output the Raman spectrum in units of pixels. The process of calibrating the spectrum from pixels to Raman shift, typically reported in wavenumbers, is a multi-step process. We recommend daily calibration of the instrument, as even subtle shifts in the instrument optics may produce an observable change in the Raman shift. It is important that the user ensures that the laser spot corresponds to the visual and spectrometer sampling point preceding sample acquisition. If necessary, beam alignment can be performed using manual or computerised beam-steers and should be checked regularly to optimise spectral acquisition. Dependent upon the instrument manufacturer, there are manual and/or automatic calibration options. The first calibration step is conversion of pixels to wavelength. This is achieved by measuring a calibration lamp, such as Neon or Argon that has multiple atomic emission lines, and using a non-linear model to fit the spectral peaks to the CCD pixel. The second and third steps are converting wavelength to wavenumber units but forming the inverse of the wavelength and correction of the laser wavelength using reference materials with defined Raman signatures. The easiest reference material to use is silicon, which has a sharp peak at 520.5 cm^{-1} and any band shifts can be easily offset[127]. Other Raman reference materials include cyclohexane, acetaminophen or Teflon. Luminescent or broad band emission standards can be used to derive a system response function measure to correct for filter, detector etaloning and quantum efficiency effects[128]. Further performance validation tests can include the measurement of a well characterised protein in order to check SNR and calibration consistency between measurements. National Institute of Standards and Technology (NIST) standards are also routinely used as calibration references.

Spectral quality. Optimising experimental parameters is an iterative process, requiring an element of trial-and-error in order to obtain a method for optimum spectral quality and high SNR. Spectral quality is governed by instrumentation, sample suitability and ultimately time constraints. Within sampling parameters, it is possible to attenuate the laser power so that optimum power levels are incident on the sample providing the highest Raman signal for weak scatterers, whilst preventing detector saturation for intense scatterers[124]. Following power adjustments, the exposure time of the laser on the sample can be increased, multiplying the intensity of scattered photons and therefore the spectral quality. The user also has the option to accumulate multiple spectra, increasing the signal intensity and reducing baseline interference. Longer exposure times and co-additions can significantly impact on sampling times, resulting in a potential trade-off between the spectral range, quality and time availability.

Spectral contamination. Cosmic rays are sporadic background artefacts recognised by sensitive detectors, that manifest in Raman spectra as narrow bandwidth spikes[129]. Most instrument software packages contain cosmic ray removal algorithms that allow the user to selectively eliminate cosmic rays, as well as algorithms in processing packages for automated cosmic ray removal [130]. Peaks from known contaminants, such as paraffin in FFPE tissues, can be removed using cutting techniques or computational algorithms[112]. For samples containing extrinsic fluorescence, photobleaching could be utilised to suppress interference from contaminants. Exposing samples to the incident light source for a few seconds prior to acquiring spectra on the detector, effectively 'bleaching' the fluorescent contaminants, has been shown to reduce oscillating baselines[131].

Experimental design: data processing

Spectroscopic studies can rapidly generate large datasets that require computational processing in order to derive biochemical information. Depending on the specific spectral acquisition and experimental objectives, it is possible to extract informative images, spectral biomarkers and patterns and also classify samples due to their spectral fingerprint. As a rule, data processing can be divided into three distinct steps: (i) dataset pre-processing, (ii) feature extraction, and/or (iii) classification[79]. **Table 2** provides a list of available analysis software. There is also a strong relationship between Raman and IR spectral analysis (for reviews of this technique, see[79,87,130,132]).

Pre-processing. Before any further analytical steps, the quality of the spectral datasets should be assessed and pre-processing applied to improve the accuracy of the study by minimising insignificant variability[130]. At this point, spectra should be corrected for cosmic rays and the quality of the spectra can be visually assessed. There are some circumstances when visual inspection of spectra show clear outliers including significant spectral contamination, fluorescence or very poor SNR[133]. In those cases, obvious outliers can be removed from the dataset. Other spectra may be outliers in the dataset which may be identified using a Q-test or other outlier detecting algorithm to reduce user bias. We recommend sagacious and limited removal of spectra from a data set because it has an impact particularly on hyperspectral images and small datasets.

Raman spectra are particularly prone to noise and data may require noise reduction to enhance spectral quality. The first approach to improve the quality of the Raman signal would be alter spectral acquisition settings, such as using increased integration times and higher laser power. Sample preparation may also be adjusted by pre-concentrating or photobleaching the sample. If these approaches are inadequate, then spectra can be computationally manipulated post-acquisition to improve the SNR. Principal component analysis (PCA) is a powerful technique in Raman pre-processing that can effectively reduce the spectra into a defined number of principal components (PCs) that account for significant spectral variance[134]. This technique can be employed to reconstruct spectra using only significant PCs, therefore retaining important spectral data whilst removing background noise[135]. Another approach is noise reduction by Savitzky-Golay (SG) or kernel smoothing, minimum noise fraction transform and wavelet de-noising (WDN) techniques that can filter high frequency noise[132]. While smoothing spectra does reduce the apparent noise, we note that these processes also degrade spectral features and recommend limited and cautious use of smoothing.

Sample fluorescence and thermal fluctuations of the CCD can significantly impact on the spectral baseline and therefore require baseline correction. Polynomial baseline fitting attempts to estimate the unknown background. This is often dependent upon user defined polynomial points, and can effectively abolish sloped or oscillatory baselines[136,137]. Care should always be taken with any baseline subtraction routines as they can introduce unintended artefacts. Alternatively, 1st or 2nd order differentiation, coupled with SG smoothing, can be applied to mathematically remove contributions from scalar offsets or baseline slopes, whilst simultaneously resolving overlapped peaks[138]. Unlike polynomial fitting which yields spectra with conventional morphologies, derivative spectra are transformed and do not have a regular Raman appearance.

Following baseline correction spectra may also require normalisation to correct for sample and experimental variables, such as thickness and density. Vector normalisation and min-max normalisation are two popular methods that can be applied to spectra following any baseline correction algorithms without significant impact on spectral features[130]. Amide I peak normalisation is also commonly utilised in IR and Raman studies[79]. However, this technique is not appropriate following differentiation, due to the shift of the typical Amide bands, or for observing protein alterations as all values are scaled to 1.

Finally, an optional data reduction step can be included in the pre-processing procedure to optimise statistical analysis. As highlighted previously, large spectral datasets can often present a significant computational burden, due to the many absorbance intensities contained in a single spectrum. By truncating the spectrum to shorter wavenumber ranges this burden can be reduced, especially if the range focuses on individual Raman peaks[139]. Alternatively, data reduction algorithms, such as partial least squares (PLS) or PCA can be employed to reduce individual spectra down to a few significant factors and has been widely implemented as both a pre-processing and feature extraction step[132].

The selection of pre-processing steps and the order in which they are conducted has been shown to have an major impact on the outcomes of spectral analysis, complicating development of a universal approach[140]. Wherever possible we advise the use of derivative baseline correction and vector normalisation, although this does require high SNR data, it can be particularly effective in diagnostic studies. In addition whenever necessary the use of noise and data reduction tools can be applied.

Feature extraction. In both exploratory and diagnostic studies, it is important to transform individual spectra into appropriate variables that confer biological information. Feature extraction methods range from very simple, such as defining a band area, to considerably more complicated computational functions. Feature extraction can broadly be split into two distinct approaches, feature construction and feature selection. These approaches are comprehensively reviewed by Trevisan *et al*[132]. Briefly, feature construction can be defined as the creation of new features in a dataset that can infer otherwise obscured information, for example, the previously mentioned linear methods PCA and PLS. This can be exceptionally important for diagnostics, biomarker extraction and pattern recognition in otherwise homogenous datasets and also play an important role in hyperspectral imaging, as individual pixels can be reduced to single values relating to spectral intensity or variance[79]. Feature selection approaches extrapolate existing features from the dataset, such as specific wavenumbers, that can be used to determine spectral biomarkers and/or feed into diagnostic

frameworks[141]. Techniques such a genetic algorithm, multivariate curve resolution and successive projection algorithm have proven particularly popular as feature extraction methods, as only informative variables are included in the resultant model[142].

Classification. Classification of samples based on their spectra is often desirable, in both imaging and diagnostic studies, as spectra can be categorised based on prior user input (supervised classification) or spectral variance alone (unsupervised). Unsupervised classification typically relies upon a clustering technique, of which hierarchical cluster analysis, k-means clustering and fuzzy C-means clustering are three popular options[80]. PCA is another unsupervised approach that is commonly used to extract key variables describing the largest variance within a dataset. In imaging studies, this approach requires no prior knowledge of the sample in question, and produces information-rich pseudo-spectral images that are ideal for exploratory studies[143]. The loadings or spectral features used to calculate the scores or weights images can provide key information on molecular distributions in a sample.

For diagnostic analysis, supervised classification is most commonly used, as the desired outputs are dependent upon class labels provided by the user. This could be gold standard histopathology, cell type, or an alternative measure of class. These class assignments are taken into consideration when implementing the classification technique; using a proportion of the dataset, referred to as a 'training dataset'. A classifier is then able to categorise a separate 'test dataset' accordingly. It is a common approach to validate the classification outputs using an independent dataset. Linear discriminant classifiers (LDC), artificial neural networks, and support vector machines (SVM) are particularly common machine learning techniques that effectively classify spectral data, although other techniques are available[144]. Semi-supervised classification is an alternative option when some, but not all, data classes can be determined[145,146].

Sample size. Classifier performance and validity is highly influenced by sample size and thus is an essential aspect of experimental design. The number of spectra within a dataset can easily surpass millions, as a sufficient number of spectra per sample and class type is required to account for the inherent sample intra- and inter-class variability. However, it must be remembered that the number of spectra is not representative of the sample size and frequently the number of independent samples is much smaller and may, for example, be determined by the overall patient number in a given study. It has been shown that a sample size of between 75 - 100 is sufficient to train a classification model with good precision and validation[147]. Cross-validation may be a suitable solution in smaller datasets as resampling approaches can repeat or iterate different training and test datasets for a defined number of times, effectively using as much of the dataset as possible[132,148]. However, full independent testing of a previously developed classification algorithm is the ideal approach, if the sample numbers permit.

MATERIALS

For all materials listed alternative suppliers can be used, unless otherwise stated.

REAGENTS

! CAUTION We note that most of the listed reagents are chemicals with potential hazards. Users should consult with the chemical's Material Safety Data Sheet and the appropriate facility safety guidelines prior to handling chemicals. Raman spectra for each reagent used may be provided by the reagent supplier or accessed via a Raman database provided by the instrument manufacturer.

- Sample acquisition: Fixed tissue, fresh tissue, cryosectioned tissue, fixed cells, live cells, biofluids and powders from a variety of biological specimens.
  ▲ CRITICAL Research carried out with human subjects must be compliant with the Helsinki Declaration. Research carried out with animals must be approved by the local institutional review board or animal use ethical board. Approvals must be established before the experiment's start.
- Fixative agents: -
  - o Formalin, 10 % (vol/vol) (Sigma-Aldrich, HT5011128)
  - o Glutaraldehyde, 25% (vol/vol) (Sigma-Aldrich, G5882)
  - o Ethanol (Fisher Scientific, E/0600DF/17)
  - o Methanol (Fisher Scientific, A456-212)
  - o Acetone (Fisher Scientific, A19-1)
  - o Osmium tetroxide, 2% (wt/vol) (Sigma-Aldrich, 75633)
- Paraffin wax pelletised with added polymers 57-58°C (Fisher Scientific, 12624077)
- De-waxing agents: -
  - o Xylenes (Sigma-Aldrich, 534056)
  - o HistoChoice® Clearing Agent (Sigma-Aldrich, H2779)
- Liquid nitrogen (BOC, CAS no. 7727-37-9)
- Isopentane (Fisher Scientific, P/1030/08)
- SERS nanoparticles:-
  - o Gold 150 nm (1.66×10^9 particles/ml) (BBI Solutions, EM. GC150)
  - o Gold 40 nm (9.00×10^{10} particles/ml) (BBI Solutions, EM. GC40)
  - o Silver 40 nm (2.6×10^9 particles/ml) (BBI Solutions, EM. SC40)
- Virkon® Disinfectant (Fisher Scientific, NC0480633)

REAGENT SETUP

- **Fixatives:** Must be diluted to appropriate concentration using phosphate buffered or saline H_2O solution. Can be refrigerated at 4°C for several weeks before use.
- **SERS nanoparticles:** Require dilution to appropriate concentration dependent upon experimental parameters.
- **Biological samples:-**
  - o **Fixed tissue.** Generally, tissues are fixed using an appropriate concentration of chemical fixative, followed by alcohol dehydration. Tissues can then be either desiccated stored at room temperature (20-22°C) for a number of months, or paraffin embedded. An embedding base mould should be filled with molten wax using a wax dispenser. The sample should be carefully placed into the wax at the desired orientation and allowed to cool. FFPE tissue blocks can be stored at room temperature indefinitely.
  - o **Fresh tissue.** *In vitro* studies, such as when tissue has been excised from the independent sample, will be prone to sample degradation and thus should be

analysed as soon after excision as possible. Refrigeration at 4°C or imitating the natural environment may be sufficient for several days of storage. Tissues may be snap-frozen using liquid nitrogen and isopentane for long term storage with minimal effect on sample integrity.

- o **Fixed cells.** Fixed using appropriate chemical fixative or a preservative buffer and can be stored at room temperature. Cells can be gentled centrifuged to form a concentrated cell pellet, wherever necessary.
- o **Live cells.** Dependent upon desired experimental aims, live cells must be kept in optimum living conditions throughout the study, including nutrient requirements via media, as well as temperature, pH, light and gas conditions.
- o **Biofluids.** Once obtained, biofluids can be stored at -80°C for several years to prevent degradation. Prior to sample preparation, biofluids should be thawed using a water bath at 35-37°C.

EQUIPMENT

! CAUTION We note that the listed equipment has potential hazards. Users should consult with the equipment instruction manual and the appropriate facility safety guidelines prior to use.

- Electronic Equipment:
  - o Raman microspectrometer (**Table. 3** illustrates a number of commercially available systems)
  - o Computer system: A standard computer should be sufficient for basic spectral acquisition and basic data analysis. As computational demands rise in correlation to data processing with increased complexity, it is advised that a system with sufficient RAM access (upwards of 4 GB in 32-bit PCs), and a high-speed processor is used.
- Microtome (Thermo Fisher Scientific, 902100A)
- Wax dispenser (Electrothermal, MH8523B)
- Sectioning bath (Electrothermal, MH8517)
- Centrifuge (Thermo Fisher Scientific, 75002410)
- Desiccator (Thermo Fisher Scientific, 5311-0250)
- Substrates:-
  - o Glass slides (Fisher Scientific, 12657956)
  - o CaF_2 slides (Crystran, CAFP10-10-1)
  - o Quartz slides (UQG Optics, FQM-2521)
  - o Gold-coated slides (Platypus Technologies, AU.0500.ASLI)
  - o Aluminium-coated slides (EMF, AL134)
  - o Quartz vial (Starna Cells, 1-Q-1)
- Accessories:-
  - o Laser power meter (Coherent, 1098293)
  - o Microtome blades (Leica Biosystems, 14035843490)
  - o Desiccant (Sigma-Aldrich, 13767)
  - o Embedding base moulds (Leica Biosystems, 38VSP58167)
  - o Magnets, weights or mounting putty

EQUIPMENT SETUP

- **Software**: **Table 3** describes the available software options that come as standard with each specific commercial instrument. These software options are essential for spectral acquisition and may also provide data processing provisions. However, for specialised spectral analysis we direct the user to **Table 2**, where alternative software options are listed.

PROCEDURE

In the following protocol we focus upon three experimental examples: (i) *in vivo* spectral exploration of live plant samples (suitable for fresh tissue analysis); (ii) Imaging of fixed human tissue; and (iii) SERS classification of endometrial cancer using biofluids.

Sample preparation

1| Prepare the samples and mount onto appropriate Raman substrates following option (A) for *in vivo* plant analysis; option (B) for FFPE tissue samples; or option (C) for SERS biofluids analysis; or option (D) for cell analysis.

(A) Plant tissue ● TIMING 5 min
- (i) Obtain plant specimen from sample community.
- (ii) Place sample upon a suitable Raman substrate (dependent upon thickness of tissue). ▲CRITICAL Leaf tissue in most plant species should be sufficiently thick to avoid background interference from the substrate. However, roots tissues for example may require a high quality substrate.
- (iii) Secure sample using a weight, adhesive, or magnetic if using a magnetic stage accessory.

(B) FFPE tissue ● TIMING 30 min + dewaxing (1.5 h)
- (i) Acquire FFPE tissue blocks from pathology laboratory with appropriate ethical approval.
- (ii) Place FFPE block upon a cool, preferably frozen surface, for at least 10 minutes. ▲CRITICAL STEP Cooling hardens the wax which therefore facilitates smooth sectioning.
- (iii) Position block in microtome sample holder and begin to trim using large sections (up to 25 μm) until the surface on the tissue is exposed.
- (iv) Alter the microtome sectioning thickness appropriately, 5-10 μm is common in histopathology, and begin to cut sample ribbons.
- (v) Float individual ribbons in a heated sectioning bath at 40 - 44°C. ▲CRITICAL STEP Gentle heat will help relax the wax and enable effective mounting. Heat must not surpass 45°C as the sample will begin to degenerate as the wax approaches its melting point.
- (vi) Carefully mount samples onto appropriate Raman substrate, by using the slide to lift the ribbon out of the water. ▲CRITICAL STEP If the sample requires dewaxing proceed to step (vii), if not proceed to step (ix).
- (vii) Allow the slide to dry for 30 minutes.

▲CRITICAL STEP The sample must be adequately adhered to the slide surface prior to dewaxing, else the sample can be lost.

(viii) Immerse the tissue in xylene for 5 minutes. Repeat twice to ensure wax is adequately removed.

(ix) Clear xylene residues using a 15 minutes 100% ethanol wash, followed by a further 15 minute wash with 90% ethanol and a final wash for 15 minutes with 70% ethanol.

(x) Store at room temperature (20-22 °C) in a dry environment until analysis.

■ PAUSE POINT Samples can be stored for up to one year, before significant degradation can occur.

(C)   SERS biofluids ● TIMING 10 min + drying (24 h)

(i) Obtain biofluids specimen with suitable ethical approval.

■ PAUSE POINT Samples can be stored at -80 °C for 1-2 years dependent upon license constraints.

(ii) Apply 200 µl of biofluids and 200 µl of colloidal nanoparticles onto appropriate Raman substrate (volumes and ratios can be altered accordingly).

▲CRITICAL STEP The user has three distinct methods for nanoparticle and sample mixture: (i) Apply nanoparticle solution to substrate, allow to dry and subsequently add sample; (ii) the reverse of the above process, where the sample is added first followed by the nanoparticles; or (iii) combine the two solutions together and apply mixture onto substrate.

(iii) Allow to dry prior to analysis. Larger sample volumes require extended drying times and thus we recommend overnight as standard for continuity.

(D) Cells ● TIMING 12 hr for cell attachment + 15 min sample prep

■ PAUSE POINT splitting cells at approximately 60% confluence greatly diminishes the number of lipid droplets per cell, which can be caused by cell "stress".

(i) Seed a known number of cells onto the appropriate Raman substrate in supplemented cell culture media, under sterile conditions, and allow to incubate overnight at 37°C.

(ii) After 12 hours or more, remove cells from incubator and aspirate off cell culture media. Wash samples with warmed sterile PBS.

(iii) For live cell imaging, samples can be imaged in a warmed sterile PBS solution or warmed sterile filtered 0.9% NaCl saline solution.

▲CRITICAL STEP Ensure all reagents are warmed to 37°C. Warming reagents to 37°C reduces the shock to the cells and helps maintain overall morphology.

(i) For cell fixation, after media has been removed, wash cells with warmed PBS.

(ii) Immerse cells in warmed 10% formalin for 10 min for fixation.

(iii) Wash thrice in warmed sterile PBS to remove any trace of the fixative.

▲CRITICAL STEP Ensure all reagents are warmed to 37°C. Warming reagents to 37°C reduces the shock to the cells and helps maintain overall morphology.

Samples can be then measured either dry or in sterile filtered 0.9% NaCl saline solution.

■ **PAUSE POINT** Samples can be stored for up to 3 months at 4°C in sterile NaCl saline solution.

Spectral acquisition

2| Switch on Raman microspectrometer and open instrument operating software.

▲**CRITICAL STEP** Carefully read the instrument operating manual and become aware of the operating features, specifications and safety operating procedures.

3| Determine suitable instrumentation options, including laser wavelength, detector type and suitable optics.

4| Calibrate and align the spectrometer using a calibration source.

5| Mount the sample upon microscope stage for spectral acquisition.

6| Using the microscope at a chosen magnification, examine the sample and focus upon a feature of interest.
? TROUBLESHOOTING
7| Determine the sampling area using the microscope and operating software using a point mapping (A) or imaging mapping (B) approach.

(A) Point mapping ● **TIMING 1 sec - 5 min (per spectrum)**
  (i)   Selectively, or randomly, chose numerous points in the sampling area to interrogate (anywhere between 5 and 50 spectra would be typical).
        ▲**CRITICAL STEP** In SERS studies, enhanced spectra are obtained from molecules in close proximity to nanoparticles. In a point map approach we advise the user isolates nanoparticle aggregations and samples from this region.
        ? TROUBLESHOOTING
(B) Image mapping ● **TIMING 1 sec - 1 min (per spectrum)**
  (i)   Select a mapping area using a 'shape fill' option (usually rectangular filled).
  (ii)   Determine the number of spectra to be acquired within the mapping area by altering the step size.
        ▲**CRITICAL STEP** Be aware of the sampling area in use (typically ~1 μm), as step sizes above this will result in under-sampling and below will result in over-sampling.

8| Input optimum spectral parameters to ensure best SNR, spectral quality within an appropriate acquisition time dependent upon the sample suitability. Determining the laser power at the sample using the designated spectral parameters may be conducted using a laser power meter.

▲**CRITICAL STEP** Point mapping is generally less time restricted due to smaller spectra numbers and thus spectral quality should be optimised. On the other hand, image mapping acquires a high number of spectra that can significantly elongate acquisition times and consequently a judgement between acquisition time and spectral quality should be made. Sample integrity must also be considered over long sampling periods. Furthermore, by

obtaining large numbers of spectra, multivariate approaches can extract much spectral detail from noisy data.

9| Acquire sample measurement.

10| Save measurements prior to data processing and convert to universal format if appropriate.

■ **PAUSE POINT** Datasets can be stored until data processing

Data pre-processing ● TIMING 10 – 60 min (dependent upon dataset size)

▲**CRITICAL STEP** Steps 11-16 highlight possible pre-processing steps that can be performed on the spectral dataset in an advised order. All these steps are optional and can be applied in multiple combinations with varying effects on visual output. We recommend that the user observes the spectral output at each stage to inspect any adverse effects on the data sets. For further information, we direct the reader to **Experimental design. data processing**.

11| Load data in selected data processing software (see **Table 2**).

12| Screen the data for anomalies and poor quality spectra using quality tests.

13| Apply a noise reduction technique on the dataset, such as PCA, to improve SNR.

14| Conduct a baseline correction algorithm to account for fluorescence interference.

15| Perform a data normalisation approach to account for confounding sample variables.

16| If necessary, perform a data reduction technique such as simple truncation or PCA to reduce the number of variables in the dataset.

Data analysis

17| Choose an appropriate data analysis approach that will extract the required information from your point spectra or images, dependent upon the desired analysis goal. The options we describe here are exploratory analysis (A), for pattern finding and biomarker extraction, and diagnostic (B) analysis. For more information, please refer to **Experimental design. Data processing.**

▲**CRITICAL STEP** All timings are estimated for typical datasets containing anywhere between 500 – 5000 spectra. Any deviations from this range will alter the timings accordingly.

(A) Exploratory ● TIMING 15 – 60 min (dependent upon dataset size)
    (i)   Input pre-processed dataset into a feature extraction algorithm, either construction or selection.
        ▲**CRITICAL STEP** For imaging, continue to step (ii); for point spectra, move to step (iv).
    (ii)   For image analysis, use feature extraction outputs to assign a scalar value to each spectrum (such as wavenumber intensities or PC score).

    (iii)   Use colour gradient, or code for scalar values.

    (iv)   Visualise analysis output for biomarker and pattern extraction.

(B)   **Diagnostic** ● **TIMING** **1 – 4 h (dependent upon dataset size)**

    ▲**CRITICAL STEP** For supervised classification, the dataset must be split into training and a *test* dataset.

    (i)   Input pre-processed dataset into a feature extraction algorithm, either construction or selection, in training mode.

    (ii)   Apply this trained feature extraction model to a *training* dataset.

    (iii)   Train a classification algorithm using this *training* dataset.

    (iv)   Input the *test* dataset into the trained FE model followed by the trained classification algorithm.

        ▲**CRITICAL STEP** Cross-validation is recommended on small sample sizes.

    (v)   Obtain a class estimation per spectrum.

    (vi)   For imaging, this class estimation can be assigned a scalar value and visualized using a colour code.

(C)   **Diagnosis (unsupervised)** ● **TIMING** **15 – 60 min (dependent upon dataset size)**

    (i)   Use a clustering classification algorithm to categorize spectra based on spectral variance.

    (ii)   Assign each cluster a numerical value/colour and visualize.

? TROUBLESHOOTING

Step 6| Maintaining optical focus

It can be difficult to maintain optical focus if the sample is not secured sufficiently, or if the surface is not adequately flat, and can have significant impact on spectral quality. To improve stability, bulky samples should be held in place using magnets or weights when possible, or secured with adhesive materials such as tape. Small amount of mounting clay/putty on the underside of hard specimens may help secure them onto a substrate. This should also help flatten some samples, but it may be necessary to physically smooth sample surfaces, such as when analysing powders. Increased stability stages are available and are recommended in systems using high magnification (100×) as small stage movements can have more significant effects at high magnification.

Step 7A(i)| Obtaining enhanced spectra

As enhancement is reliant upon direct sampling of an area in close proximity to a nanoparticle, the probability of acquiring enhanced spectra is particularly low. We recommend isolating nanoparticle aggregates, which unlike monomers, may be visible under magnification. A good knowledge of the sample in question is needed, and we advise carefully studying subtle differences between control and SERS samples. The use of larger nanoparticles (>100 nm) may also aid nanoparticle aggregate identification [62]. Otherwise, using an automatic mapping procedure may produce enhanced spectra without the need of significant user input.

Step 9| Insufficient spectral quality

As data is obtained, the user can identify four key indicators of spectral quality; (i) low SNR, (ii) fluorescent baseline, (iii) saturation of the CCD and (iv) photoablation.

(i) Low SNR - Indicated by a strong background signal in comparison to Raman peaks, is indicative of insufficient Raman signal as a consequence of the sample or the experimental parameters. First, check that the system is appropriately aligned, that optical focus is optimised and the sample is illuminated by the laser. If possible, increase the laser power, followed by the exposure time and then an increased number of co-additions. If this is not possible, consider system alterations including an alternative laser wavelength or Raman approach that is better suited to your sample (**Table 1**).

(ii) Fluorescence baseline - A range of fluorescence manifestation can be observed dependent upon the excitation wavelength in use, but is principally indicated by a strong, broad and featureless spectrum, similar to the fluorescence wavelength profile (Fig. 7A). Fluorescence is predominantly due to sample characteristics although can be influenced by background conditions, such as lights. We advise utilising a UV or NIR laser wavelength, or implementing an alternative Raman approach that can overcome fluorescence, such as SERS or SRS. Photobleaching may also reduce fluorescence contributions, with due caution of photoablation.

(iii) Saturation of the CCD - Each CCD will have a limit to how many Raman scattered photons that it can measure, and when exceeded this can be observed by a feature such as a flat line effect on the spectrum (Fig. 7B). As saturation is indicative of increased Raman scattering, this issue is usually found in strong scattering materials, samples of a large volume and in SERS studies. This can be overcome by attenuating the laser power and/or exposure time upon the sample. Saturation cannot be corrected in post-acquisition processing.

(iv) Photoablation - Some samples will be sensitive to high laser powers and thus can burn once exposed. Sometimes a burned sample will result in a saturated CCD and some residue may be on the objective. If a burned sample is observed, inspect the microscope objective for residue and the CCD signal for saturation. When this occurs visible damage to the sample can occur and increased intensity of amorphous carbon peaks. To avoid photodamage, the laser power and exposure times should be decreased accordingly, or a different laser wavelength could be used. If these approaches are inadequate, the laser power can be attenuated at the sample surface by using a larger laser spot, achieved by using a lower magnification setting or altering the laser focus.

Step 12 | Spectral contaminants

Cosmic rays or interference from sample contaminants can render spectra unusable unless they are removed. Cosmic rays can be identified as sharp spikes with narrow bandwidths that can often overshadow true Raman peaks (Fig. 7C). These should be removed using cosmic ray removal algorithms that are generally available in instrument operation software. Raman peaks associated with spectral contaminants such as paraffin (Fig. 7D), can occur due to inadequate sample preparation; however these can also be removed with respective computational algorithms, or simply ignored if studying spectral regions unaffected by their presence.

● TIMING

We provide approximate times for each of the steps. All timing will be dependent upon the size of the dataset as sample processing and computational analysis are both proportional to the number of samples employed.

Step 1A: 5 min

Step 1B: 30 min - 2 h

Step 1C: 10 min - 24 h

Step 7A: 1 -5 min per spectrum (average 15 spectra per sample ~30 min)

Step 7B: 1 sec - 1 min per spectrum (average 1000 spectra per sample ~3 hours)

Steps 11-16: 10 min - 1 h

Step 17A: 15 min - 1 h

Step 17B: 1 - 4 h

Step 17C: 15 min - 1 h

Minimum experimental time of ~1 hour per sample

Maximum experimental time of ~32 hours per sample

Anticipated Results

Pre-processing options

Figure 5 is an example of the effects of pre-processing on spectra with strong fluorescence baseline interference (785 nm laser). Spectra were truncated to a defined spectral region (1700 - 700 cm^{-1}) to account for Raman peaks present in this particular sample. A WDN step was conducted to smooth the appearance of the data and thus reduce unwanted noise. With 1st derivative baseline correction, the SG smoothing function is run simultaneously to account for loss of spectral quality using this approach. This example illustrates how effectively differentiation can account for baseline fluorescence, however spectral noise is increased and interpretability is reduced due to the transformation of the spectrum. On the other hand, polynomial baseline correction results in spectra with conventional appearance and sufficiently relieved of background interference. Vector normalisation was used in both incidences to attribute for confounding sample features, such as thickness; the effect of this is a slight reduction in variance between classes.

Classification of blood plasma and serum using SERS

Table 4 illustrates the effect of pre-processing, feature extraction and classification approaches on SERS of blood plasma and serum in endometrial cancer patients[62]. The classification rate was calculated by averaging the sensitivity (% of cancer spectra correctly defined as 'cancer') and specificity (% of control spectra correctly defined as 'control'). The

purpose of this study was to determine if blood serum or plasma were significantly better at identifying endometrial cancer and whether surface enhancement contributed to improved enhancement. From this study, plasma samples produced improved classification rates compared to blood serum which can be associated with the inclusion of clotting proteins. A non-specific SERS labelling approach was used, which in this instance did not contribute to increased classification compared to the spontaneous Raman approach (control). The effect of pre-processing can be seen when comparing the classification rates in serum and plasma, with the former performing better with polynomial baseline correction and the latter performing best with differentiation correction. Generally, the classification approach used had a small effect on rates of classification, although this may not be the case in other circumstances. A variety of pre-processing and classification approaches should be employed and compared, in order to assess their diagnostic efficiency.

Mapping of endometrial tissue

Figure 6 depicts FFPE endometrial tissue in control patients in order to identify the epithelial lining of the endometrial crypts. These tissue structures seen throughout the three examples, are found ubiquitously throughout endometrial tissue initiating at the lumen and spiralling toward the myometrium (not visible). Each map was acquired using an increasing exposure time (1→3), ultimately increasing the spectral quality at each pixel. The Raman data was analysed using the multivariate technique PCA, with PC1 and PC2 effectively identifying the connective tissue (red) and the epithelial layer surrounding the crypt lumen despite alterations in exposure times.

Acknowledgements

Competing Financial Interests

There are no competing financial interests.

Author Contribution Statements

F.L.M. is the principal investigator who conceived the idea for and finalised the manuscript; H.J.B. wrote and compiled the manuscript and figures. B.B., G.C., K. E.-W., L.A., M.J.W., M.R.M. and N.S. provided information throughout the article, whilst N.J.F. and P.L.M.-H., provided feedback on the manuscript.

Legends to figures

Figure 1. Schematic identifying light scattering following laser exposure upon a sample surface. Photons of light are focused upon the sample through the microscope objective at a defined magnification. When they interact with chemical bonds within the biological specimen, electrons are excited to virtual energy levels. These biological molecules return to the original energy level by emitting a photon of light, known as elastic or Rayleigh scattering, or it can undergo an energy shift and return at a lower (Stokes) or higher (anti-Stokes) energy levels, known as Raman scattering. Fluorescence can occur when electrons are excited to electronic energy levels and return to the ground energy level by emitting a photon of light at a longer wavelength.

Figure 2. Raman spectroscopy workflow diagram highlighting key aspects of experimental design including instrumentation, sample preparation, spectral acquisition and data processing with representative examples. It is not possible to conduct some options in combination, such as an InGaAs detector with UV wavelengths, and thus experimental design requires thorough planning prior to experimentation. Options provided are not exhaustive, and other alternatives are available.

Figure 3. Generalised overview of instrumentation options within a typical spontaneous Raman spectroscopic microscope system. The beam path will be vary slightly between manufacturers and thus may contain additional optical components.

Figure 4. Simplified overview of the effect of laser excitation wavelength on fluorescence background. Visible lasers are largely susceptible to background due to the defined wavelength profile of fluorescence (black), whereas high energy UV (purple) and lower energy IR (red) wavelengths are relatively free from fluorescence.

Figure 5. A brief overview of pre-processing options in Raman spectral data analysis and their contribution to spectral transformation. Wavelet de-noising is conducted as a smoothing process, prior to polynomial baseline correction and vector normalisation. 1st order differentiation and Savitzky-Golay smoothing are conducted simultaneously to account for noise introduction in the spectra.

Figure 6. Raman map of fixed endometrial tissue focusing upon uterine glands that spiral throughout the tissue. Samples are FFPE and were mounted upon BaF_2 slides prior to dewaxing. **(A)** images are white light images of the tissues, whilst **(B)** displays PCA maps. Epithelial tissue is depicted in yellow, whereas connective tissue is shown in red. Spectra were all obtained using a 785 nm laser with 25 mW of power at the sample, a 1200 l/mm grating and a step size of 1 μm. **(1B)** was acquired with 5 seconds exposure time and is comprised of 47 × 48 pixels; **(2B)** 10 seconds, with 74 × 81 pixels; and **(3B)** 15 seconds, with 55 × 53 pixels.

Figure 7. Examples of common troubleshooting issues during spectral acquisition. **(A)** Shows typical fluorescence from a plant sample using a 785 nm laser; **(B)** saturation of a CCD detector; **(C)** cosmic ray at ~1220 cm^{-1} and; **(D)** paraffin contamination from an embedded biological sample.

Figure 1.

Figure 2.

Figure 3.

EXCITATION SOURCE

Beam Expander

Rayleigh Filters

Focusing Mirror

Grating

DETECTOR

Focusing Mirror

Sample

Microscope

Figure 4.

Figure 5.

RAW RAMAN SPECTRA

Wavenumber (cm^{-1})

WAVELET DENOISE

POLYNOMIAL BASELINE CORRECTION

1st ORDER DIFFERENTIATION

VECTOR NORMALISATION

VECTOR NORMALISATION
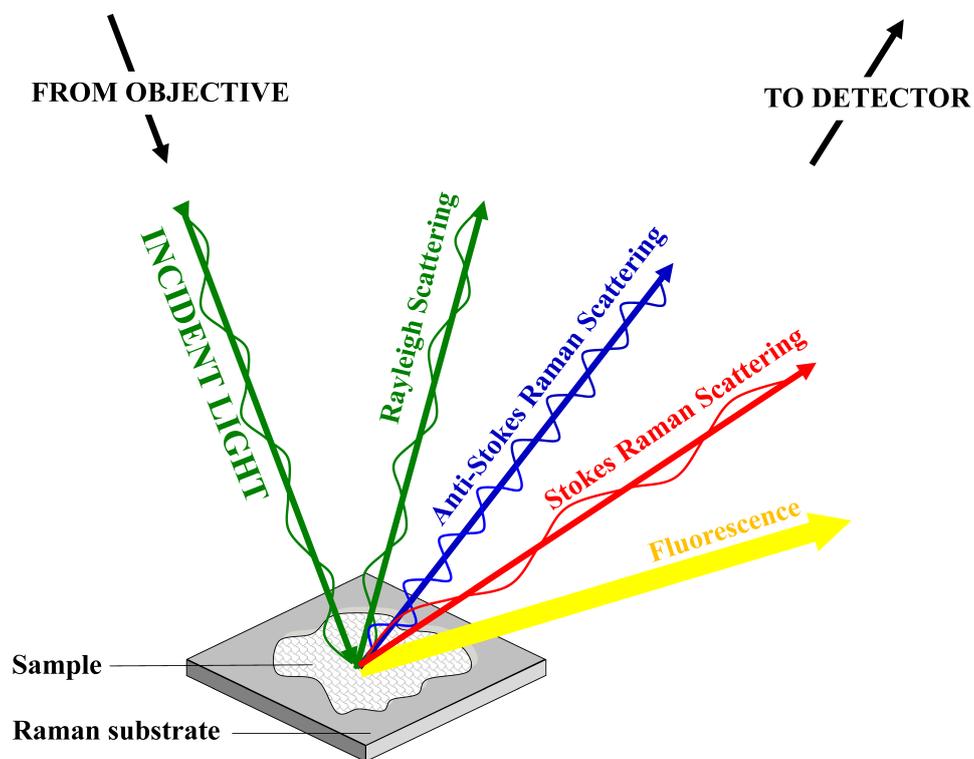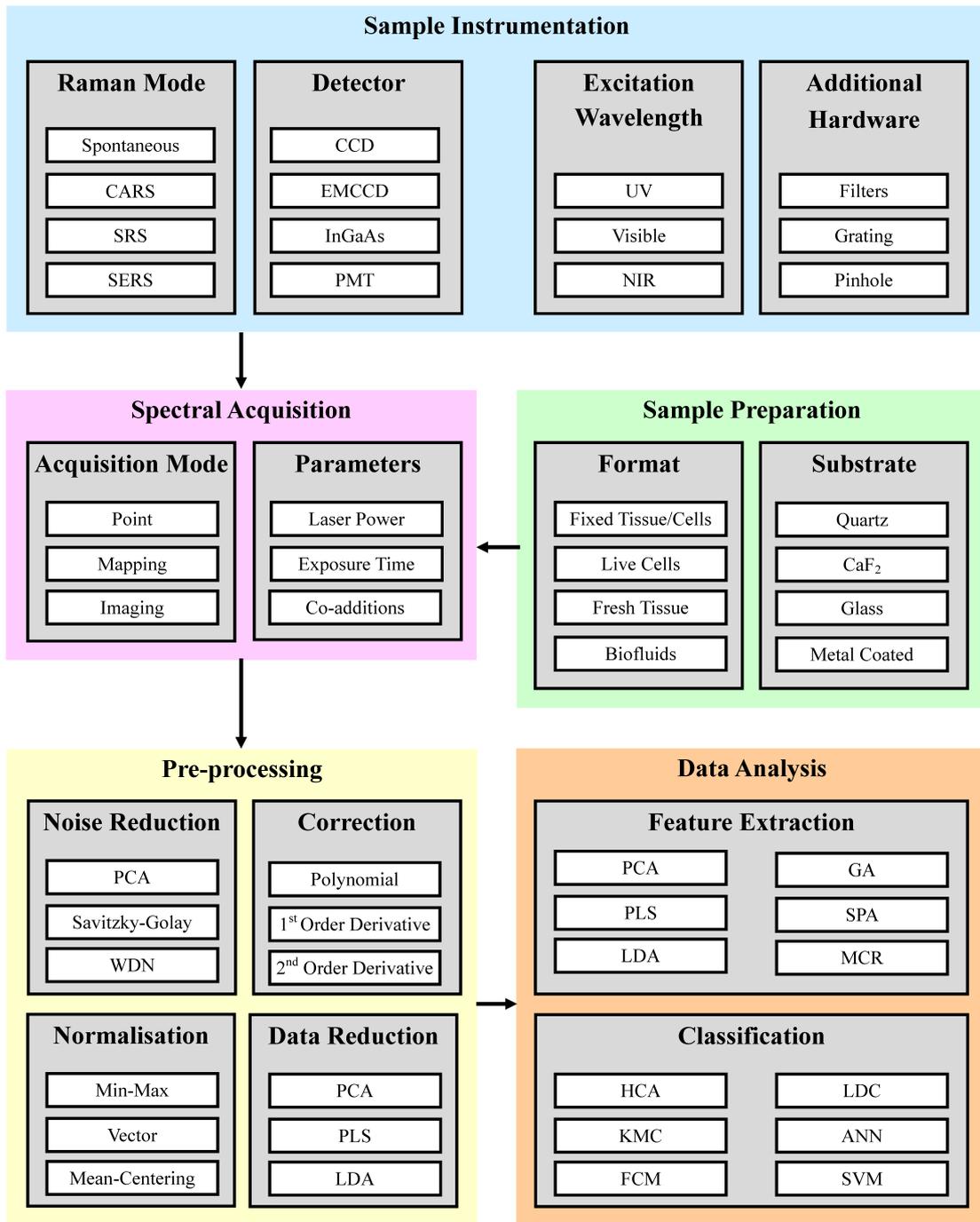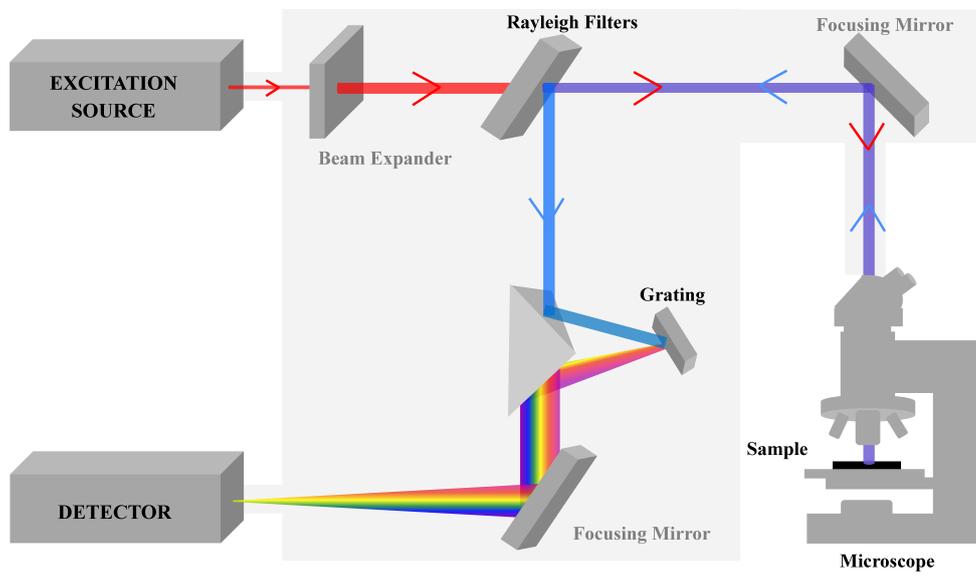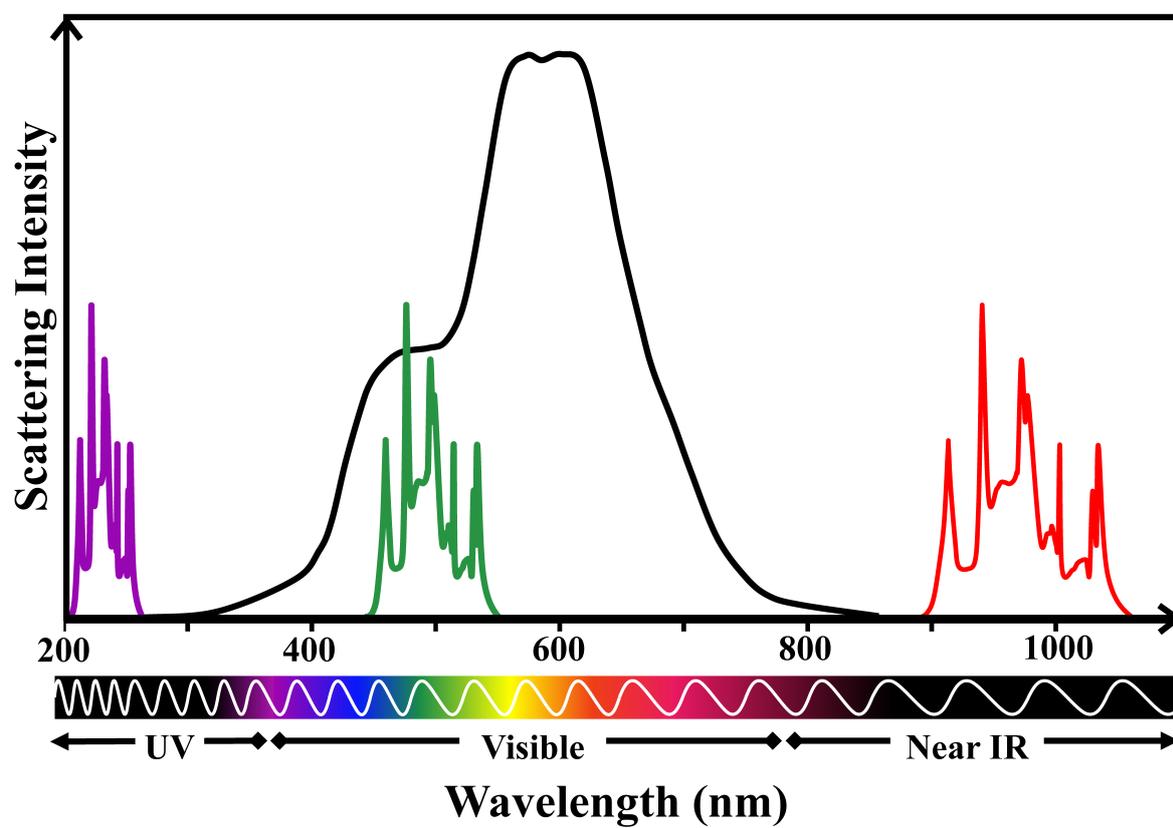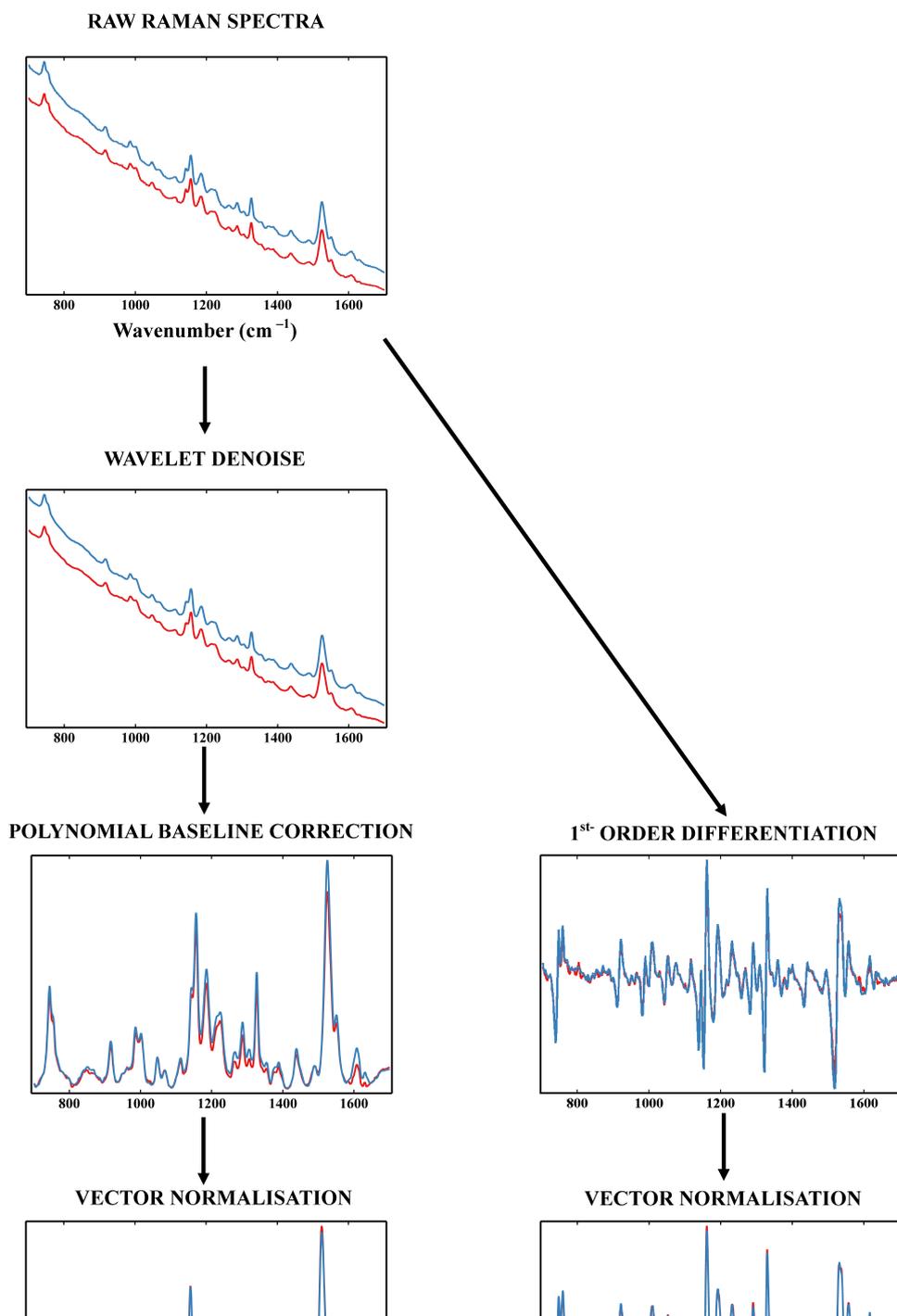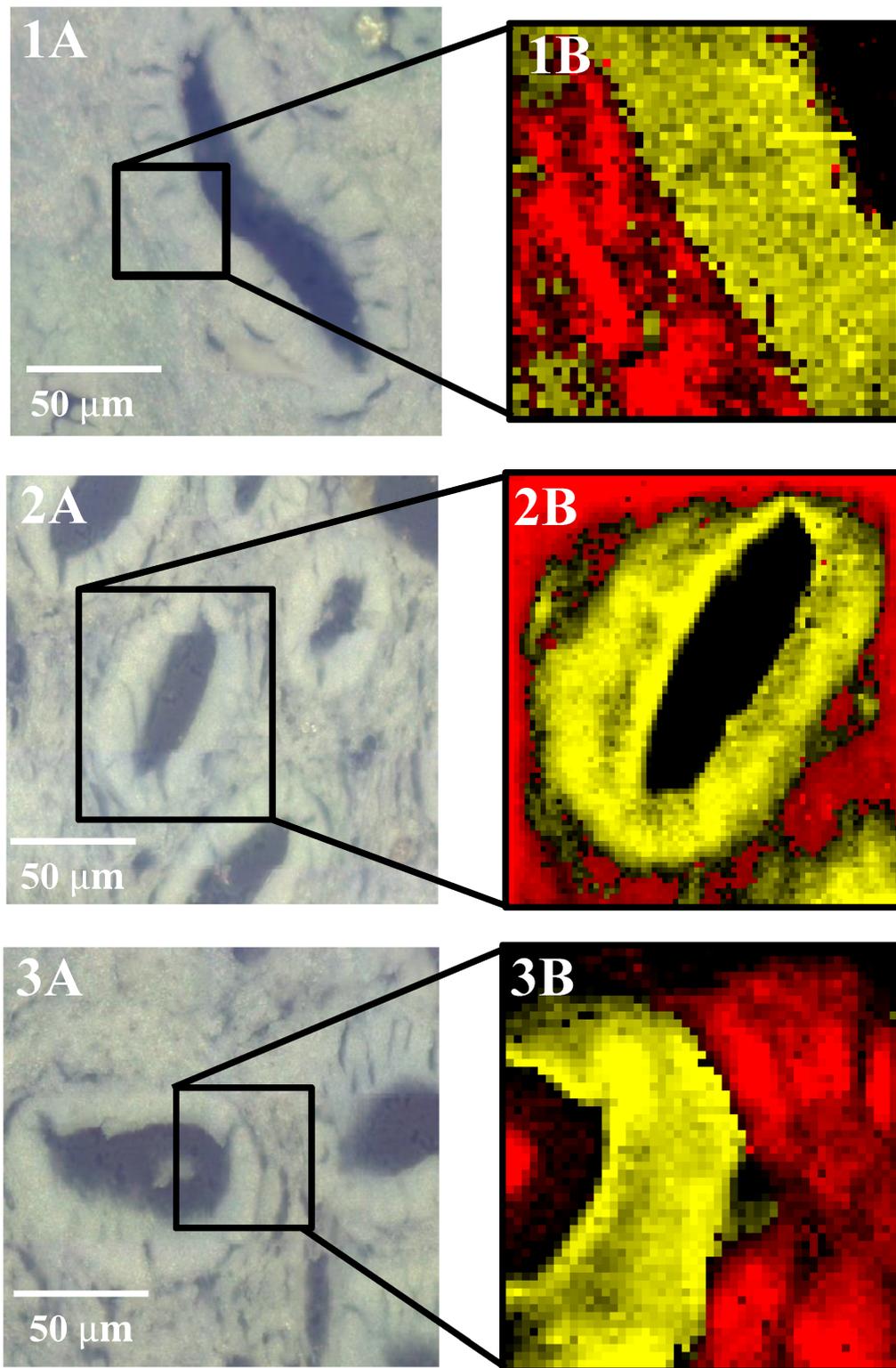
Figure 6.

Figure 7.
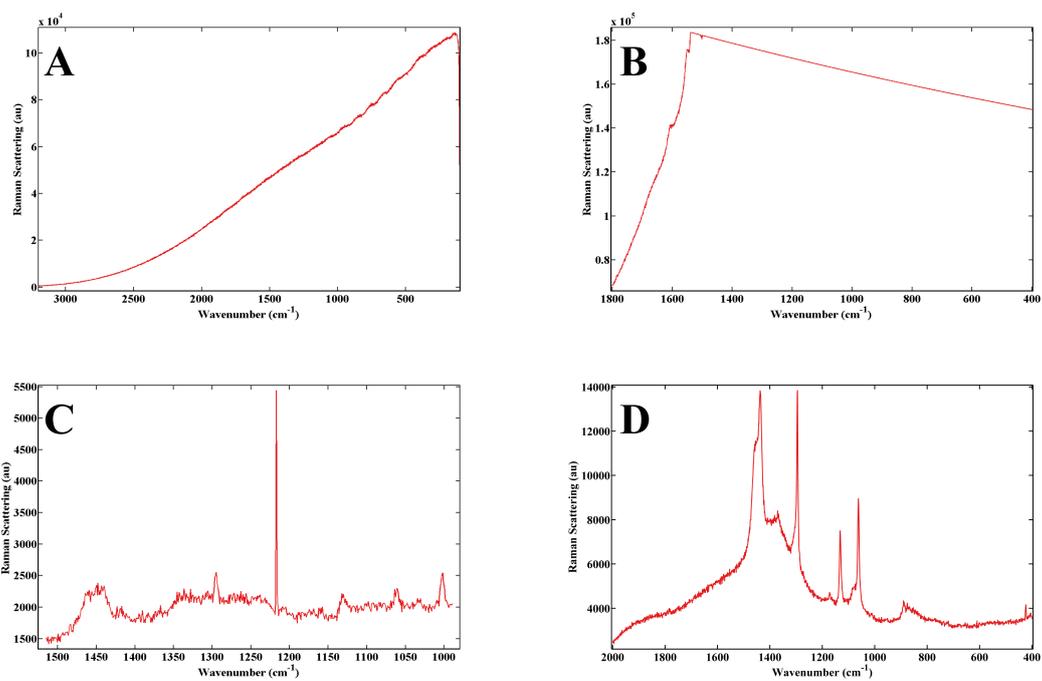
Table 1. Raman spectroscopy method variations and the respective experimental benefits and limitations.

Technique	Adaption	Benefits	Limitations	Applications/References
Coherent anti-Stokes Raman scattering (CARS)	Non-linear approach that utilises multiple laser frequencies; a pump (ω_p) and Stokes (ω_s) field that combine, tuned to a frequency equivalent to a specific molecular vibration, thus generating a strong anti-Stokes signal ($\omega_{as} = 2\omega_p\text{-}\omega_s$) that detects vibrational coherence	10^3-10^6 increased signal[149] High sensitivity 3-D imaging	Nonresonant background can dominate weak resonant signals	Cell imaging[150,151] Tissue imaging[152,153] Cancer diagnosis[154,155] Pharmaceuticals[156,157]
Confocal Raman microscopy	Addition of a confocal microscope that allows depth measurements within tissue. A pinhole is used within the spectrometer to reject stray light, effectively reducing the collimation of the beam and information to be derived from an alternative focal plane	High sensitivity High lateral and depth resolution 3-D imaging Rapid acquisition	Diffraction-limited resolution	Cell imaging[158,159] Tissue imaging[160] Cancer diagnosis[161-163] Pharmaceuticals[164,165] Plant cell imaging[74,75]
Drop coating deposition Raman spectroscopy (DCDRS)	Sample preparation for fluid analysis that drops a small volume of sample onto a flat substrate and allows to dry	Accurate preparation of biofluids Small volumes of fluid required (2 – 10 µl)	Not completely free from 'coffee ring' effect	Biofluid analysis[119] Protein quantification[117,166]
`FT-Raman	Non-dispersive system that utilises Fourier transformation using a Michaelson interferometer	High throughput High resolution Free from fluorescence	Low scattering intensity Limited to IR measurements Detector noise limited	Pharmaceuticals[13,167] Plant materials[71,168]
Kerr-gated Raman spectroscopy	A linear technique using repeated laser pulses and a Kerr gate that capture Raman light temporally (up to 3 picoseconds)	Depth measurements several millimetres Fluorescence rejection High sensitivity	Not completely free from fluorescence Better performance when in conjunction with SERDS	Depth profiling in human tissue[169,170,171]
Polarised Raman spectroscopy (PRS)	Polarised light with a specific electric field vector will only obtain spectral information from specific vibrational modes dependent on their orientation in relation to the incident beam	Information regarding molecular structure and orientation	Not applicable to most samples Lose spectral information Time consuming	Collagen orientation[172] Plant photosystems[173]
Raman Optical Activity (ROA)	Using right- and left-circularly polarised incident light, the small changes in Raman scattering can be indicative of optical activity of discrete molecular vibrations.	Structural information from specific conformations of chiral molecules	Circular intensity differences are very small Vibrational coupling in ROA signals can prevent accurate band assignment	Biopolymer analysis[174-176] Pharmaceuticals[177,178]
Resonance Raman spectroscopy (RRS)	Exploits the 'resonance effect' observed when the laser frequency matches (or approaches) that of an electronic transition of the sample or compound in question	Up to 6 orders of magnitude increase in signal[179]	Susceptible to fluorescence interference	Plant photosystems[179] Human biology[54,55,180,181]
Shifted excitation Raman difference spectroscopy (SERDS)	Nonlinear approach that obtains two spectra at marginally different laser frequencies and creates a difference spectrum by subtracting the two, thus removing background fluorescence	Fluorescence rejection Increased sensitivity	Difference spectra reconstructed using peak fitting Prone to error	Live cell[182] Animal tissue[183] Human tissue[184]

Spatially offset Raman spectroscopy (SORS)	Continuous, low intensity laser beams are employed to illuminate the surface of the sample and Raman spectra are then derived at distinct distances away from this point. A scaled subtraction between these spectra reveals alterations indicative of the underlying sub-surface layers	Depth measurements up to several millimetres	Relatively weak signal	Cancer diagnosis[32,185,186] Chemical analysis beneath physical obstructions[187,188]
Surface enhanced spatially offset Raman spectroscopy (SESORS)	A combination of SERS and SORS approaches, able to detect SERS nanoparticles introduced into turbid samples	Detect SERS signals up to 50 mm beneath the sample surface[189]	Requires nanoparticle introduction	Depth measurements[190-193]
Stimulated Raman scattering (SRS)	Nonlinear approach using a pump (ω_p) and Stokes (ω_s) field that are tuned to a defined frequency representative of a molecular vibrations (ω_p- ω_s). When this occurs ω_p observes a stimulated Raman loss (SRL) in energy, and ω_s observes a stimulated Raman gain (SRG). The transferred intensity is proportional to the biochemical constituents	Not affected by fluorescence and nonresonant background High sensitivity (1 in 10^6 photons) High spatial resolution	Prone to interference from strong Raman scatterers Limited to measurement of one Raman peak per acquisition	Cell imaging[58] Plant imaging[23,194,195]
Surface enhanced Raman scattering (SERS)	Due to surface plasmon resonance, a metal surface with nanoscale roughness can significantly increase the electric field when excited by a laser. Thus when adsorbed to a biomolecule, these nanoparticles result in greatly enhanced Raman scattering	$10^3 – 10^{10}$ enhancement Below diffraction limit resolution Quench fluorescence Low detection limit Molecular labelling	Lack of reproducibility Can reduce band intensity of high frequency modes Molecular selectivity to nanoparticle adherence	Single molecule detection[196,197] Tumour targeting[198] Live cell analysis[199] Pharmaceuticals[200] Cancer diagnosis[62,201,202] Bacteria identification[16,203] Plant materials[204]
Surface enhanced resonance Raman scattering (SERRS)	Combination of SERS and RRS, where the laser frequency	Up to 10^{15} enhancement[205] Cumulative benefits of both SERS and RRS	Increased experimental complexity	Biomolecule detection[206] Protein analysis[207,208]
Tip enhanced Raman spectroscopy (TERS)	Based upon the same electromagnetic and chemical theory as SERS, TERS utilises an atomic force microscope (AFM) tip coated with SERS active metal. When placed into close proximity to the sample, it results in enhanced scattering	Tip dependent spatial resolution. Low detection limit Quench fluorescence Below diffraction limit resolution	Increased experimental complexity Sample heating at tip apex	Microbiology[18] Biochemical imaging[209]
Total Internal Reflection (TIR) Raman spectroscopy	The sample is placed in contact with a reflective prism, through which a laser beam is reflected, producing an evanescent wave that penetrates the sample below	Defined penetration depth	Reduced surface sensitivity	Plant materials[68]
Transmission Raman	Raman scattered light is captured on the opposite side of laser illumination	Depth measurements up to 30 mm Suitable for opaque materials	Interference from surface molecules	Cancer diagnosis[210] Pharmaceuticals[211,212]

Table 2. Data analysis software for Raman spectral datasets

Software	Website	License
CytoSpec	http://www.cytospec.com/ftir.php	Commercial
ImageLab	http://www.imagelab.at/en_home.html	Commercial
Matlab® Biodata Toolbox EMSC IRootLab MIA MCR-ALS PLS Raman Processing Program	http://www.mathworks.com/matlabcentral/fileexchange/22068-biodata-toolbox http://www.models.life.ku.dk/emsctoolbox https://code.google.com/p/irootlab/ http://www.eigenvector.com/software/mia_toolbox.htm http://www.cid.csic.es/homes/rtaqam/tmp/WEB_MCR/welcome.htm http://www.eigenvector.com/software/pls_toolbox.htm http://cares.wayne.edu/rp/	Commercial Open source Open source Open source Commercial Open source Commercial Open source
Origin for Spectroscopy	http://www.originlab.com/index.aspx?go=Solutions/Applications/Spectroscopy	Commercial
PeakFit®	http://www.sigmaplot.com/products/peakfit/peakfit.php	Commercial
Python PyChem PyVib2	 http://pychem.sourceforge.net/ http://pyvib2.sourceforge.net/	Open source Open source Open source
R hyperSpec	 http://hyperspec.r-forge.r-project.org/	Open source
The Unscrambler® X	http://www.camo.com/	Commercial

Table 3. Commercially available instruments and corresponding operational software.

Manufacturer	Instruments		Software
BaySpec	Agility™	S	Spec 20/20
	RamSpec™	S	
	RamSpec™ - HR	S	
	Nomadic™	M	
	MovingLab™	MP	
Bruker Optics	SENTERRA	M	OPUS
	RamanScopeIII	M	
	MultiRAM	S	
	RAM II	S	
	BRAVO	SP	
Horiba Scientific	XploRA™ Series	M	LabSpec 6
	LabRAM HR Evolution	M	
	Triple Raman Spectrometers	S	
	OEM Miniature Raman	SP	
JASCO	NRS-5000 Series Raman	M	Spectra Manager
	NRS-7000 Series Raman	M	II™
	RMP-10	SP	
Kaiser Optical Systems	RamanRXN1™	M	iC Raman™
	Raman WorkStation™	M	HoloMap™
Ocean Optics	Maya2000 Pro	S	OceanView
	QE Pro	S	
	Ventana Series	S	
Ondax	THz-Raman® Spectroscopy Systems	M	N/A
	TR-MICRO	M	
Renishaw	InVia confocal Raman microscope	M	WiRE
	RA100 portable Raman Analyser	SP	
SciAps	Inspector300	S	NuSpec
	Inspector500	S	NuSpec Pro
	ReporteR	SP	
Thermo Fisher Scientific	DXR™ Series	M	ValPro™ System
	FirstDefender™ RM Chemical	SP	Qualification
	Identification System	SP	
	FirstDefender™ RMX Handheld	SP	
	Chemical Identification	SP	
WITec	Alpha300 Series	M	WITec Suite
	apyron ∞	SM	

S – Spectrometer, **M** – Microspectrometer, **P** – Portable

Table 4. Classification rates (% classification ± standard deviation) of control and endometrial cancer patients from blood serum and plasma samples, using surface enhanced Raman spectroscopy (SERS) with two different diameter gold nanoparticles and spontaneous Raman ('No SERS'). Two pre-processing approaches and two classification approaches are used to illustrate varied performances. Polynomial baseline correction was conducted with a polynomial order of 5 and first order Savitzky-Golay (SG) differentiation used 9 filter co-efficients with a polynomial order of two. A support vector machine (SVM; optimised C and ɣ parameters) classifier was implemented without a feature extraction step, whereas principal component analysis (PCA; optimised number of PCs) was employed prior to a linear discriminant classifier (LDC). The cells are colour coded dependent upon the classification rate (yellow to red).

Sample	Pre-processing	SERS	Classification Approach (%)	
			SVM	PCA-LDC
Serum	Polynomial → Vector Normalisation	No SERS	89.58 ± 07.18	88.56 ± 11.35
		40 nm	69.42 ± 08.37	59.69 ± 19.24
		150 nm	72.69 ± 11.73	72.99 ± 14.40
	1st Order Differentiation → Vector Normalisation	No SERS	79.63 ± 09.23	79.79 ± 11.46
		40 nm	65.74 ± 09.40	65.83 ± 26.17
		150 nm	69.52 ± 17.79	64.76 ± 15.86
Plasma	Polynomial → Vector Normalisation	No SERS	97.28 ± 03.87	93.55 ± 08.18
		40 nm	79.30 ± 11.10	79.79 ± 14.27
		150 nm	73.11 ± 14.19	74.76 ± 19.81
	1st Order Differentiation → Vector Normalisation	No SERS	98.50 ± 02.69	95.07 ± 05.15
		40 nm	80.02 ± 10.39	79.11 ± 13.77
		150 nm	76.14 ± 19.14	79.81 ± 17.63

REFERENCES

1   Clemens, G., Hands, J. R., Dorling, K. M. & Baker, M. J. Vibrational spectroscopic methods for cytology and cellular research. *Analyst* **139**, 4411-4444, (2014).

2   Ellis, D. I., Cowcher, D. P., Ashton, L., O'Hagan, S. & Goodacre, R. Illuminating disease and enlightening biomedicine: Raman spectroscopy as a diagnostic tool. *Analyst* **138**, 3871-3884, (2013).

3   Downes, A. & Elfick, A. Raman spectroscopy and related techniques in biomedicine. *Sensors* **10**, 1871-1889, (2010).

4   Krafft, C. & Popp, J. The many facets of Raman spectroscopy for biomedical analysis. *Anal Bioanal Chem* **407**, 699-717, (2015).

5   Movasaghi, Z., Rehman, S. & Rehman, I. U. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **42**, 493-541, (2007).

6   Chase, B. A new generation of Raman instrumentation. *Appl. Spectrosc.* **48**, 14-19, (1994).

7   Zhang, Y., Hong, H. & Cai, W. Imaging with Raman spectroscopy. *Curr. Pharm. Biotechno.* **11**, 654-661, (2010).

8   Kneipp, K., Kneipp, H., Itzkan, I., Dasari, R. R. & Feld, M. S. Surface-enhanced Raman scattering and biophysics. *J. Phys: Condens. Mat.* **14**, R597-R624, (2002).

9   Hartschuh, A., Sánchez, E. J., Xie, X. S. & Novotny, L. High-resolution near-field Raman microscopy of single-walled carbon nanotubes. *Phys. Rev. Lett.* **90**, 095503, (2003).

10   Cooper, J. B. Chemometric analysis of Raman spectroscopic data for process control applications. *Chemometr. Intell. Lab.* **46**, 231-247, (1999).

11   Widjaja, E., Zheng, W. & Huang, Z. Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines. *Int. J. Oncol.* **32**, 653-662, (2008).

12   Krafft, C., Steiner, G., Beleites, C. & Salzer, R. Disease recognition by infrared and Raman spectroscopy. *J. Biophotonics* **2**, 13-28, (2009).

13   Vankeirsbilck, T. *et al.* Applications of Raman spectroscopy in pharmaceutical analysis. *TrAC.* **21**, 869-877, (2002).

14   Ekins, S. & Sasic, S. *Pharmaceutical applications of Raman spectroscopy*. (John Wiley & Sons, 2008).

15   Frosch, T., Yan, D. & Popp, J. r. Ultrasensitive fiber enhanced UV resonance Raman sensing of drugs. *Anal. Chem.* **85**, 6264-6271, (2013).

16   Jarvis, R. M. & Goodacre, R. Discrimination of bacteria using surface-enhanced Raman spectroscopy. *Anal. Chem.* **76**, 40-47, (2004).

17   Maquelin, K. *et al.* Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* **51**, 255-271, (2002).

18   Neugebauer, U. *et al.* On the way to nanometer-sized information of the bacterial surface by tip-enhanced Raman spectroscopy. *ChemPhysChem* **7**, 1428-1430, (2006).

19   Pahlow, S., Meisel, S., Cialla-May, D., Weber, K. & Röschac, J. P. P. Isolation and identification of bacteria by means of Raman spectroscopy. *Adv. Drug Deliv. Rev.* **Forthcoming**, (2015).

20   Ashton, L., Lau, K., Winder, C. L. & Goodacre, R. Raman spectroscopy: lighting up the future of microbial identification. *Future Microbio.* **6**, 991-997, (2011).

21   Owen, C. A. *et al.* In vitro toxicology evaluation of pharmaceuticals using Raman micro-spectroscopy. *J. Cell. Biochem.* **99**, 178-186, (2006).

22   Knief, P. *et al.* Raman spectroscopy–a potential platform for the rapid measurement of carbon nanotube-induced cytotoxicity. *Analyst* **134**, 1182-1191, (2009).

23   Mansfield, J. C. *et al.* Label-free chemically specific imaging in planta with stimulated Raman scattering microscopy. *Anal. Chem.* **85**, 5055-5063, (2013).

24   Baranska, M., Roman, M., Schulz, H. & Baranski, R. Recent Advances in Raman Analysis of Plants: Alkaloids, Carotenoids, and Polyacetylenes. *Curr. Anal. Chem.* **9**, 108-127, (2013).

25   Gierlinger, N. & Schwanninger, M. The potential of Raman microscopy and Raman imaging in plant research. *Spectrosc.- Int. J.* **21**, 69-89, (2007).

26   Kallaway, C. *et al.* Advances in the clinical application of Raman spectroscopy for cancer diagnostics. *Photodiagnosis Photodyn. Ther.* **10**, 207-219, (2013).

27   Byrne, H. J. *et al.* Spectropathology for the next generation: Quo vadis? *Analyst* **140**, 2066-2073, (2015).

28   Meyer, T. *et al.* Nonlinear microscopy, infrared, and Raman microspectroscopy for brain tumor analysis. *J. Biomed. Opt.* **16**, 021113(021111)-021113(021110), (2011).

29   Fullwood, L. M. *et al.* Investigating the use of Raman and immersion Raman spectroscopy for spectral histopathology of metastatic brain cancer and primary sites of origin. *Anal. Methods* **6**, 3948-3961, (2014).

30   Gajjar, K. *et al.* Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis. *Anal. Methods* **5**, 89-102, (2013).

31   Krafft, C., Neudert, L., Simat, T. & Salzer, R. Near infrared Raman spectra of human brain lipids. *Spectrochim. Acta A* **61**, 1529-1535, (2005).

32   Stone, N., Baker, R., Rogers, K., Parker, A. W. & Matousek, P. Subsurface probing of calcifications with spatially offset Raman spectroscopy (SORS): future possibilities for the diagnosis of breast cancer. *Analyst* **132**, 899-905, (2007).

33   Crow, P., Uff, J., Farmer, J., Wright, M. & Stone, N. The use of Raman spectroscopy to identify and characterize transitional cell carcinoma in vitro. *BJU. Int.* **93**, 1232-1236, (2004).

34   Lin, D. *et al.* Colorectal cancer detection by gold nanoparticle based surface-enhanced Raman spectroscopy of blood serum and statistical analysis. *Opt. Express* **19**, 13565-13577, (2011).

35   Dekker, E. & Fockens, P. Advances in colonic imaging: new endoscopic imaging methods. *Euro. J. Gastroen. Hepat.* **17**, 803-808, (2005).

36   Stone, N., Stavroulaki, P., Kendall, C., Birchall, M. & Barr, H. Raman spectroscopy for early detection of laryngeal malignancy: preliminary results. *Laryngoscope* **110**, 1756-1763, (2000).

37   Huang, Z. *et al.* Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *Int. J. Cancer* **107**, 1047-1052, (2003).

38   Horsnell, J. D. *et al.* Raman spectroscopy–a potential new method for the intra-operative assessment of axillary lymph nodes. *Surgeon* **10**, 123-127, (2012).

39   Lloyd, G. R. *et al.* Discrimination between benign, primary and secondary malignancies in lymph nodes from the head and neck utilising Raman spectroscopy and multivariate analysis. *Analyst* **138**, 3900-3908, (2013).

40   Stone, N., Kendall, C., Shepherd, N., Crow, P. & Barr, H. Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *J. Raman Spectrosc.* **33**, 564-573, (2002).

41   Kendall, C. *et al.* Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. *J. Pathol.* **200**, 602-609, (2003).

42   Bergholt, M. S. *et al.* In vivo diagnosis of esophageal cancer using image-guided Raman endoscopy and biomolecular modeling. *Technol. Cancer Res. T.* **10**, 103-112, (2011).

43   Crow, P. *et al.* Assessment of fiberoptic near-infrared Raman spectroscopy for diagnosis of bladder and prostate cancer. *Urology* **65**, 1126-1130, (2005).

44   Patel, I. I. & Martin, F. L. Discrimination of zone-specific spectral signatures in normal human prostate using Raman spectroscopy. *Analyst* **135**, 3060-3069, (2010).

45   Patel, I. I. *et al.* Segregation of human prostate tissues classified high-risk (UK) versus low-risk (India) for adenocarcinoma using Fourier-transform infrared or Raman microspectroscopy coupled with discriminant analysis. *Anal Bioanal Chem* **401**, 969-982, (2011).

46   Crow, P. *et al.* The use of Raman spectroscopy to identify and grade prostatic adenocarcinoma in vitro. *Br. J. Cancer* **89**, 106-108, (0000).

47   Rashid, N. *et al.* Raman microspectroscopy for the early detection of pre-malignant changes in cervical tissue. *Exp. Mol. Path.* **97**, 554-564, (2014).

48   Patel, I. I. *et al.* High contrast images of uterine tissue derived using Raman microspectroscopy with the empty modelling approach of multivariate curve resolution-alternating least squares. *Analyst* **136**, 4950-4959, (2011).

49   Krishna, C. M. *et al.* Raman spectroscopy studies for diagnosis of cancers in human uterine cervix. *Vib. Spectrosc.* **41**, 136-141, (2006).

50   Lyng, F. M. *et al.* Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool. *Exp. Mol. Pathol.* **82**, 121-129, (2007).

51   Mitchell, A. L., Gajjar, K. B., Theophilou, G., Martin, F. L. & Martin-Hirsch, P. L. Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting. *J. Biophotonics.* **7**, 153-165, (2014).

52   Wood, B. R. *et al.* Raman imaging of hemozoin within the food vacuole of Plasmodium falciparum trophozoites. *FEBS Lett.* **554**, 247-252, (2003).

53   Wood, B. R. *et al.* Tip-enhanced Raman scattering (TERS) from hemozoin crystals within a sectioned erythrocyte. *Nano Lett.* **11**, 1868-1873, (2011).

54   Wood, B. R. & McNaughton, D. Resonance Raman spectroscopy in malaria research. *Expert Rev. Proteomic.* **3**, 525-544, (2006).

55   Wood, B. R. *et al.* Resonance Raman spectroscopy reveals new insight into the electronic structure of β-hematin and malaria pigment. *J. Am. Chem. Soc.* **126**, 9233-9239, (2004).

56   Bonnier, F. *et al.* Imaging live cells grown on a three dimensional collagen matrix using Raman microspectroscopy. *Analyst* **135**, 3169-3177, (2010).

57   Notingher, I. & Hench, L. L. Raman microspectroscopy: a noninvasive tool for studies of individual living cells in vitro. *Expert Rev. Med. Devices* **3**, 215-234, (2006).

58   Zhang, X. *et al.* Label-free live-cell imaging of nucleic acids using stimulated Raman scattering microscopy. *ChemPhysChem* **13**, 1054-1059, (2012).

59   Kneipp, J., Kneipp, H., Rajadurai, A., Redmond, R. W. & Kneipp, K. Optical probing and imaging of live cells using SERS labels. *J. Raman Spectrosc.* **40**, 1-5, (2009).

60   Farhane, Z., Bonnier, F., Casey, A. & Byrne, H. Raman micro spectroscopy for in vitro drug screening: subcellular localisation and interactions of Doxorubicin. *Analyst* **140**, 4212-4223, (2015).

61   Meister, K. *et al.* Label-free imaging of metal–carbonyl complexes in live cells by Raman microspectroscopy. *Angew. Chem. Int.* **49**, 3310-3312, (2010).

62   Butler, H. J. *et al.* Gold nanoparticles as a substrate in bio-analytical near-infrared surface-enhanced Raman spectroscopy. *Analyst* **140**, 3090-3097, (2015).

63   Ackermann, K. R., Henkel, T. & Popp, J. Quantitative Online Detection of Low-Concentrated Drugs via a SERS Microfluidic System. *ChemPhysChem* **8**, 2665-2670, (2007).

64   Harper, M. M., Dougan, J. A., Shand, N. C., Graham, D. & Faulds, K. Detection of SERS active labelled DNA based on surface affinity to silver nanoparticles. *Analyst* **137**, 2063-2068, (2012).

65   Barrett, L., Dougan, J. A., Faulds, K. & Graham, D. Stable dye-labelled oligonucleotide-nanoparticle conjugates for nucleic acid detection. *Nanoscale* **3**, 3221-3227, (2011).

66   Faulds, K., Smith, W., Graham, D. & Lacey, R. Assessment of silver and gold substrates for the detection of amphetamine sulfate by surface enhanced Raman scattering (SERS). *Analyst* **127**, 282-286, (2002).

67   Gill, D., Kilponen, R. & Rimai, L. Resonance Raman scattering of laser radiation by vibrational modes of carotenoid pigment molecules in intact plant tissues. *Nature* **227**, 743-744, (1970).

68   Greene, P. R. & Bain, C. D. Total internal reflection Raman spectroscopy of barley leaf epicuticular waxes in vivo. *Colloid. Surface. B* **45**, 174-180, (2005).

69   Schulz, H. & Baranska, M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib. Spectrosc.* **43**, 13-25, (2007).

70   Sene, C., McCann, M. C., Wilson, R. H. & Grinter, R. Fourier-transform Raman and Fourier-transform infrared spectroscopy (an investigation of five higher plant cell walls and their components). *Plant Physiol.* **106**, 1623-1631, (1994).

71   Schulz, H., Baranska, M. & Baranski, R. Potential of NIR-FT-Raman spectroscopy in natural carotenoid analysis. *Biopolymers* **77**, 212-221, (2005).

72   Rösch, P., Popp, J. & Kiefer, W. Raman and surface enhanced Raman spectroscopic investigation on Lamiaceae plants. *J. Mol. Struct.* **480–481**, 121-124, (1999).

73   Schmidt, M. *et al.* Raman imaging of cell wall polymers in Arabidopsis thaliana. *Biochem. Biophys. Res. Comm.* **395**, 521-523, (2010).

74   Gierlinger, N. & Schwanninger, M. Chemical imaging of poplar wood cell walls by confocal Raman microscopy. *Plant Physiol.* **140**, 1246-1254, (2006).

75   Gierlinger, N., Keplinger, T. & Harrington, M. Imaging of plant cell walls by confocal Raman microscopy. *Nat. Protoc.* **7**, 1694-1708, (2012).

76   Asher, S. A. & Johnson, C. R. Raman spectroscopy of a coal liquid shows that fluorescence interference is minimized with ultraviolet excitation. *Science (New York, N.Y.)* **225**, 311-313, (1984).

77   Butler, H. J., McAinsh, M. R., Adams, S. & Martin, F. L. Application of vibrational spectroscopy techniques to non-destructively monitor plant health and development. *Anal. Methods* **7**, 4059-4070, (2015).

78   Baranski, R., Baranska, M. & Schulz, H. Changes in carotenoid content and distribution in living plant tissue can be observed and mapped in situ using NIR-FT-Raman spectroscopy. *Planta* **222**, 448-457, (2005).

79   Baker, M. J. *et al.* Using Fourier transform IR spectroscopy to analyze biological materials. *Nat. Protoc.* **9**, 1771-1791, (2014).

80   Diem, M., Romeo, M., Boydston-White, S., Miljković, M. & Matthäus, C. A decade of vibrational micro-spectroscopy of human cells and tissue (1994–2004). *Analyst* **129**, 880-885, (2004).

81   Dumas, P., Sockalingum, G. D. & Sule-Suso, J. Adding synchrotron radiation to infrared microspectroscopy: what's new in biomedical applications? *Trends Biotechnol.* **25**, 40-44, (2007).

82   Bhargava, R. Infrared spectroscopic imaging: the next generation. *Appl. Spectrosc.* **66**, 1091-1120, (2012).

83   Palonpon, A. F. *et al.* Raman and SERS microscopy for molecular imaging of live cells. *Nat. Protoc.* **8**, 677-692, (2013).

84   Felten, J. *et al.* Vibrational spectroscopic image analysis of biological material using multivariate curve resolution–alternating least squares (MCR-ALS). *Nat. Protoc.* **10**, 217-240, (2015).

85   Kong, L. *et al.* Characterization of bacterial spore germination using phase-contrast and fluorescence microscopy, Raman spectroscopy and optical tweezers. *Nat. Protoc.* **6**, 625-639, (2011).

86   Li, J. F. *et al.* Surface analysis using shell-isolated nanoparticle-enhanced Raman spectroscopy. *Nat. Protoc.* **8**, 52-65, (2013).

87   Martin, F. L. *et al.* Distinguishing cell types or populations based on the computational analysis of their infrared spectra. *Nat. Protoc.* **5**, 1748-1760, (2010).

88   Angel, S., Carrabba, M. & Cooney, T. The utilization of diode lasers for Raman spectroscopy. *Spectrochim. Acta A* **51**, 1779-1799, (1995).

89   Müller, A. *et al.* Diode laser based light sources for biomedical applications. *Laser Photon. Rev.* **7**, 605-627, (2013).

90   Yan, F. & Vo-Dinh, T. Surface-enhanced Raman scattering detection of chemical and biological agents using a portable Raman integrated tunable sensor. *Sensor. Actuat. B-Chem.* **121**, 61-66, (2007).

91   Moore, D. & Scharff, R. J. Portable Raman explosives detection. *Anal Bioanal Chem* **393**, 1571-1578, (2009).

92   Draux, F. *et al.* Raman spectral imaging of single living cancer cells: a preliminary study. *Analyst* **134**, 542-548, (2009).

93   Creely, C., Volpe, G., Singh, G., Soler, M. & Petrov, D. Raman imaging of floating cells. *Optics express* **13**, 6105-6110, (2005).

94   Swain, R. & Stevens, M. Raman microspectroscopy for non-invasive biochemical analysis of single cells. *Biochem. Soc. T.* **35**, 544-550, (2007).

95   Zoladek, A., Pascut, F. C., Patel, P. & Notingher, I. Non-invasive time-course imaging of apoptotic cells by confocal Raman micro-spectroscopy. *J. Raman Spectrosc.* **42**, 251-258, (2011).

96   Vo-Dinh, T. *Biomedical Photonics Handbook: Biomedical Diagnostics*. Vol. 2 (CRC press, 2014).

97   Asher, S. A., Ludwig, M. & Johnson, C. R. UV resonance Raman excitation profiles of the aromatic amino acids. *J. Am. Chem. Soc.* **108**, 3186-3197, (1986).

98   Kumamoto, Y., Taguchi, A., Smith, N. I. & Kawata, S. Deep ultraviolet resonant Raman imaging of a cell. *J. Biomed. Opt.* **17**, 076001(076001)-076001(076004), (2012).

99   Ivanda, M. & Furić, K. Line focusing in micro-Raman spectroscopy. *Appl. Opt.* **31**, 6371-6375, (1992).

100   Okada, M. *et al.* Label-free Raman observation of cytochrome c dynamics during apoptosis. *PNAS. USA* **109**, 28-32, (2012).

101   Schlücker, S., Schaeberle, M. D., Huffman, S. W. & Levin, I. W. Raman microspectroscopy: a comparison of point, line, and wide-field imaging methodologies. *Anal. Chem.* **75**, 4312-4318, (2003).

102   Minamikawa, T. *et al.* Label-free detection of peripheral nerve tissues against adjacent tissues by spontaneous Raman microspectroscopy. *Histochem. Cell Biol.* **139**, 181-193, (2013).

103   Markwort, L., Kip, B., Da Silva, E. & Roussel, B. Raman imaging of heterogeneous polymers: a comparison of global versus point illumination. *Appl. Spectrosc.* **49**, 1411-1430, (1995).

104   Ma, J. & Ben-Amotz, D. Rapid micro-Raman imaging using fiber-bundle image compression. *Appl. Spectrosc.* **51**, 1845-1848, (1997).

105   Cooper, J. *et al.* Raman spectroscopy with a low-cost imaging CCD array. *Spectrochim. Acta A* **50**, 567-575, (1994).

106   LaPlant, F. in *Emerging Raman applications and techniques in biomedical and pharmaceutical fields* Vol. 1 (eds Pavel Matousek & Michael D. Morris) 1-24 (Springer, 2010).

107   Dieing, T. & Hollricher, O. High-resolution, high-speed confocal Raman imaging. *Vib. Spectrosc.* **48**, 22-27, (2008).

108   Harnly, J. M. & Fields, R. E. Solid-state array detectors for analytical spectrometry. *Appl. Spectrosc.* **51**, 334A, (1997).

109   Li, Z., Deen, M. J., Kumar, S. & Selvaganapathy, P. R. Raman Spectroscopy for In-Line Water Quality Monitoring—Instrumentation and Potential. *Sensors* **14**, 17275-17303, (2014).

110   Carriere, J. T. & Havermeyer, F. in *SPIE BiOS.* 821905 (International Society for Optics and Photonics).

111   Pitt, G. *et al.* in *Science, Measurement and Technology, IEE Proceedings-.* 241-318 (IET).

112   Tfayli, A. *et al.* Digital dewaxing of Raman signals: discrimination between nevi and melanoma spectra obtained from paraffin-embedded skin biopsies. *Appl. Spectrosc.* **63**, 564-570, (2009).

113   Ali, S. M. *et al.* Raman spectroscopic analysis of human skin tissue sections ex-vivo: evaluation of the effects of tissue processing and dewaxing. *J. Biomed. Opts* **18**, 61202, (2013).

114   Mariani, M. M., Lampen, P., Popp, J., Wood, B. R. & Deckert, V. Impact of fixation on in vitro cell culture lines monitored with Raman spectroscopy. *Analyst* **134**, 1154-1161, (2009).

115   Deegan, R. D. *et al.* Capillary flow as the cause of ring stains from dried liquid drops. *Nature* **389**, 827-829, (1997).

116   Filik, J. & Stone, N. Analysis of human tear fluid by Raman spectroscopy. *Anal. Chim. Acta* **616**, 177-184, (2008).

117   Filik, J. & Stone, N. Investigation into the protein composition of human tear fluid using centrifugal filters and drop coating deposition Raman spectroscopy. *J. Raman Spectrosc.* **40**, 218-224, (2009).

118   Bonnier, F., Petitjean, F., Baker, M. J. & Byrne, H. J. Improved protocols for vibrational spectroscopic analysis of body fluids. *J. Biophotonics* **7**, 167-179, (2014).

119   Esmonde-White, K. A., Esmonde-White, F. W., Morris, M. D. & Roessler, B. J. Characterization of biofluids prepared by sessile drop formation. *Analyst* **139**, 2734-2741, (2014).

120   Esmonde-White, K. A., Le Clair, S. V., Roessler, B. J. & Morris, M. D. Effect of Conformation and Drop Properties on Surface-Enhanced Raman Spectroscopy of Dried Biopolymer Drops. *Applied spectroscopy* **62**, 503-511, (2008).

121   Byrne, H. J., Sockalingum, G. & Stone, N. in *Biomedical applications of synchrotron infrared microspectroscopy: a practical approach*   (ed David Moss) Ch. 4, 105-142 (Royal Society of Chemistry, 2011).

122   Fullwood, L. M. *et al.* Effect of substrate choice and tissue type on tissue preparation for spectral histopathology by Raman microspectroscopy. *Analyst* **139**, 446-454, (2014).

123   Wehbe, K., Filik, J., Frogley, M. D. & Cinque, G. The effect of optical substrates on micro-FTIR analysis of single mammalian cells. *Anal Bioanal Chem* **405**, 1311-1324, (2013).

124   Lewis, I. R. & Edwards, H. *Handbook of Raman spectroscopy: from the research laboratory to the process line*. (CRC Press, 2001).

125   Liu, Z., Zhao, C., Han, L. & Mo, Y. Study on the configuration and applications of high spectral resolution Raman spectrometer. *Guang Pu Xue Yu Guang Pu Fen Xi* **30**, 567-570, (2010).

126   Wieboldt, D. Understanding Raman Spectrometer Parameters. (2010).

127   Wiberley, S. E., Colthup, N. B. & Daly, L. H. *Introduction to infrared and Raman spectroscopy*. 3rd edn, (Elsevier, 2012).

128   Stone, N., Kendall, C., Smith, J., Crow, P. & Barr, H. Raman spectroscopy for identification of epithelial cancers. *Faraday Discuss.* **126**, 141-157, (2004).

129   Zhang, L. & Henson, M. J. A practical algorithm to remove cosmic spikes in Raman imaging data for pharmaceutical applications. *Appl. Spectrosc.* **61**, 1015-1020, (2007).

130   Lasch, P. Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging. *Chemometr. Intell. Lab.* **117**, 100-114, (2013).

131   Barman, I., Kong, C.-R., Singh, G. P. & Dasari, R. R. Effect of photobleaching on calibration model development in biological Raman spectroscopy. *J. Biomed. Opt.* **16**, 011004, (2011).

132   Trevisan, J., Angelov, P. P., Carmichael, P. L., Scott, A. D. & Martin, F. L. Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives. *Analyst* **137**, 3202-3215, (2012).

133   Naumann, D. in *Biomedical Optical Spectroscopy, Proceedings of SPIE.*   (International Society for Optics and Photonics).

134   Jolliffe, I. *Principal component analysis*. (Wiley Online Library, 2002).

135   Chen, G. & Shen-En, Q. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. *IEEE Geosci. Remote Sens. Mag.* **49**, 973-980, (2011).

136   Lieber, C. A. & Mahadevan-Jansen, A. Automated method for subtraction of fluorescence from biological Raman spectra. *Appl. Spectrosc.* **57**, 1363-1367, (2003).

137   Mazet, V., Carteret, C., Brie, D., Idier, J. & Humbert, B. Background removal from spectra by designing and minimising a non-quadratic cost function. *Chemometr. Intell. Lab.* **76**, 121-133, (2005).

138   Savitzky, A. & Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627-1639, (1964).

139   Bocklitz, T., Walter, A., Hartmann, K., Rösch, P. & Popp, J. How to pre-process Raman spectra for reliable and stable models? *Anal. Chim. Acta* **704**, 47-56, (2011).

140   Heraud, P., Wood, B. R., Beardall, J. & McNaughton, D. Effects of pre-processing of Raman spectra on in vivo classification of nutrient status of microalgal cells. *J. Chemometr.* **20**, 193-197, (2006).

141   Trevisan, J. *et al.* Measuring similarity and improving stability in biomarker identification methods applied to Fourier-transform infrared (FTIR) spectroscopy. *J. Biophotonics* **7**, 254-265, (2014).

142   de Sousa Marques, A., de Melo, M. C. N., Cidral, T. A. & de Lima, K. M. G. Feature selection strategies for identification of Staphylococcus aureus recovered in blood cultures using FT-IR spectroscopy successive projections algorithm for variable selection: A case study. *J. Microbiol. Meth.* **98**, 26-30, (2014).

143   Lasch, P., Haensch, W., Naumann, D. & Diem, M. Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis. *BBA- Mol. Basis Dis.* **1688**, 176-186, (2004).

144   Ellis, D. I. & Goodacre, R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst* **131**, 875-885, (2006).

145   Lloyd, G. R. *et al.* Utilising non-consensus pathology measurements to improve the diagnosis of oesophageal cancer using a Raman spectroscopic probe. *Analyst* **139**, 381-388, (2014).

146   Balabin, R. M., Safieva, R. Z. & Lomakina, E. I. Near-infrared (NIR) spectroscopy for motor oil classification: From discriminant analysis to support vector machines. *Microchem. J.* **98**, 121-128, (2011).

147   Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C. & Popp, J. Sample size planning for classification models. *Anal. Chim. Acta* **760**, 25-33, (2013).

148   Esbensen, K. H. & Geladi, P. Principles of proper validation: use and abuse of re-sampling for validation. *J. Chemometr.* **24**, 168-187, (2010).

149   Antonio, K. A. & Schultz, Z. D. Advances in biomedical Raman microscopy. *Anal. Chem.* **86**, 30-46, (2013).

150   Parekh, S. H., Lee, Y. J., Aamer, K. A. & Cicerone, M. T. Label-free cellular imaging by broadband coherent anti-Stokes Raman scattering microscopy. *Biophys. J.* **99**, 2695-2704, (2010).

151   Zumbusch, A., Holtom, G. R. & Xie, X. S. Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering. *Phys. Rev. Lett.* **82**, 4142-4145, (1999).

152   Krafft, C., Dietzek, B. & Popp, J. Raman and CARS microspectroscopy of cells and tissues. *Analyst* **134**, 1046-1057, (2009).

153   Evans, C. L. *et al.* Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy. *PNAS. USA* **102**, 16807-16812, (2005).

154   Le, T. T., Huff, T. B. & Cheng, J.-X. Coherent anti-Stokes Raman scattering imaging of lipids in cancer metastasis. *BMC Cancer* **9**, 42, (2009).

155   Gao, L. *et al.* Label-free high-resolution imaging of prostate glands and cavernous nerves using coherent anti-Stokes Raman scattering microscopy. *Biomed. Opt. Express* **2**, 915-926, (2011).

156   Garrett, N. *et al.* Label-free imaging of polymeric nanomedicines using coherent anti-stokes Raman scattering microscopy. *J. Raman Spectrosc.* **43**, 681-688, (2012).

157   Windbergs, M. *et al.* Chemical imaging of oral solid dosage forms and changes upon dissolution using coherent anti-Stokes Raman scattering microscopy. *Anal. Chem.* **81**, 2085-2091, (2009).

158   Schuster, K. C., Reese, I., Urlaub, E., Gapes, J. R. & Lendl, B. Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy. *Anal. Chem.* **72**, 5529-5534, (2000).

159   Majzner, K. *et al.* 3D confocal Raman imaging of endothelial cells and vascular wall: perspectives in analytical spectroscopy of biomedical research. *Analyst* **138**, 603-610, (2012).

160   Caspers, P. J., Lucassen, G. W. & Puppels, G. J. Combined in vivo confocal Raman spectroscopy and confocal microscopy of human skin. *Biophys. J.* **85**, 572-580, (2003).

161   Choi, J. *et al.* Direct observation of spectral differences between normal and basal cell carcinoma (BCC) tissues using confocal Raman microscopy. *Biopolymers* **77**, 264-272, (2005).

162   Yu, C., Gestl, E., Eckert, K., Allara, D. & Irudayaraj, J. Characterization of human breast epithelial cells by confocal Raman microspectroscopy. *Cancer Detect. Prev.* **30**, 515-522, (2006).

163   Haka, A. S. *et al.* Identifying microcalcifications in benign and malignant breast lesions by probing differences in their chemical composition using Raman spectroscopy. *Cancer Res.* **62**, 5375-5380, (2002).

164   Breitenbach, J., Schrof, W. & Neumann, J. Confocal Raman-spectroscopy: analytical approach to solid dispersions and mapping of drugs. *Pharm. Res.* **16**, 1109-1113, (1999).

165   Franzen, L., Selzer, D., Fluhr, J. W., Schaefer, U. F. & Windbergs, M. Towards drug quantification in human skin with confocal Raman microscopy. *Euro. J. Pharm. Biopharm.* **84**, 437-444, (2013).

166   Esmonde-White, K. A., Le Clair, S. V., Roessler, B. J. & Morris, M. D. Effect of Conformation and Drop Properties on Surface-Enhanced Raman Spectroscopy of Dried Biopolymer Drops. *Appl. Spectrosc.* **62**, 503-511, (2008).

167   Skoulika, S. G. & Georgiou, C. A. Rapid quantitative determination of ciprofloxacin in pharmaceuticals by use of solid-state FT-Raman spectroscopy. *Appl. Spectrosc.* **55**, 1259-1265, (2001).

168   Edwards, H., Farwell, D. & Webster, D. FT Raman microscopy of untreated natural plant fibres. *Spectrochim. Acta A* **53**, 2383-2392, (1997).

169   Morris, M. D. *et al.* Kerr-gated time-resolved Raman spectroscopy of equine cortical bone tissue. *J. Biomed. Opt.* **10**, 14014(14011)-14014(14017), (2005).

170   Prieto, M. C. H. *et al.* Use of picosecond Kerr-gated Raman spectroscopy to suppress signals from both surface and deep layers in bladder and prostate tissue. *J. Biomed. Opt.* **10**, 44006(44001)-44006(44006), (2005).

171   LouiseáRonayne, K. & WilliamáParker, A. Depth profiling of calcifications in breast tissue using picosecond Kerr-gated Raman spectroscopy. *Analyst* **132**, 48-53, (2007).

172   Galvis, L., Dunlop, J. W. C., Duda, G., Fratzl, P. & Masic, A. Polarized Raman anisotropic response of collagen in tendon: towards 3D orientation mapping of collagen in tissues. *PLoS. One* **8**, e63518, (2013).

173   Brose, K., Zouni, A., Broser, M., Müh, F. & Maultzsch, J. Polarised Raman measurements on the core complex of crystallised photosystem II. *Phys. Status Solidi* **246**, 2813-2816, (2009).

174   Blanch, E. W. *et al.* Is polyproline II helix the killer conformation? a raman optical activity study of the amyloidogenic prefibrillar intermediate of human lysozyme. *J. Mol. Biol.* **301**, 553-563, (2000).

175   McColl, I. H. *et al.* A new perspective on β-sheet structures using vibrational Raman optical activity: from poly(l-lysine) to the prion protein. *J. Am. Chem. Soc.* **125**, 10019-10026, (2003).

176   Blanch, E. W., Hecht, L. & Barron, L. D. Vibrational Raman optical activity of proteins, nucleic acids, and viruses. *Methods* **29**, 196-209, (2003).

177   Nieto-Ortega, B. *et al.* Raman optical activity spectra and conformational elucidation of chiral drugs. The case of the antiangiogenic aeroplysinin-1. *J. Phys. Chem. A* **115**, 2752-2755, (2011).

178   Yamamoto, S., Watarai, H. & Bouř, P. Monitoring the backbone conformation of valinomycin by Raman optical activity. *ChemPhysChem* **12**, 1509-1518, (2011).

179   Robert, B. Resonance Raman spectroscopy. *Photosynth. Res.* **101**, 147-155, (2009).

180   Wood, B. R., Caspers, P., Puppels, G. J., Pandiancherri, S. & McNaughton, D. Resonance Raman spectroscopy of red blood cells using near-infrared laser excitation. *Anal. Bioanal. Chem.* **387**, 1691-1703, (2007).

181   Darvin, M. *et al.* Non-invasive in vivo detection of the carotenoid antioxidant substance lycopene in the human skin using the resonance Raman spectroscopy. *Laser Phys. Lett.* **3**, 460, (2006).

182   Xie, C. & Li, Y.-q. Confocal micro-Raman spectroscopy of single biological cells using optical trapping and shifted excitation difference techniques. *J. Appl. Phys.* **93**, 2982-2986, (2003).

183   Sowoidnich, K. & Kronfeldt, H.-D. Fluorescence rejection by shifted excitation Raman difference spectroscopy at multiple wavelengths for the investigation of biological samples. *Int. Sch. Res. Net.* **2012**, 1-11, (2012).

184   Martins, M. A. d. S. *et al.* Shifted-excitation Raman difference spectroscopy for in vitro and in vivo biological samples analysis. *Biomed. Opt. Express* **1**, 617-626, (2010).

185   Matousek, P. & Stone, N. Prospects for the diagnosis of breast cancer by noninvasive probing of calcifications using transmission Raman spectroscopy. *J. Biomed. Opt.* **12**, 024008(024001)-024008(024008), (2007).

186   Keller, M. D. *et al.* Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation. *J. Biomed. Opt.* **16**, 077006-077006-077008, (2011).

187   Matousek, P. *et al.* Noninvasive Raman spectroscopy of human tissue in vivo. *Appl. Spectrosc.* **60**, 758-763, (2006).

188   Olds, W. J. *et al.* Spatially offset Raman spectroscopy (SORS) for the analysis and detection of packaged pharmaceuticals and concealed drugs. *Forensic Sci. Int.* **212**, 69-77, (2011).

189   Stone, N. *et al.* Surface enhanced spatially offset Raman spectroscopic (SESORS) imaging–the next dimension. *Chem. Sci.* **2**, 776-780, (2011).

190   Ma, K. *et al.* In vivo, transcutaneous glucose sensing using surface-enhanced spatially offset Raman spectroscopy: multiple rats, improved hypoglycemic accuracy, low incident power, and continuous monitoring for greater than 17 days. *Anal. Chem.* **83**, 9146-9152, (2011).

191   Yuen, J. M., Shah, N. C., Walsh Jr, J. T., Glucksberg, M. R. & Van Duyne, R. P. Transcutaneous glucose sensing by surface-enhanced spatially offset Raman spectroscopy in a rat model. *Anal. Chem.* **82**, 8382-8385, (2010).

192   Xie, H. n. *et al.* Tracking bisphosphonates through a 20 mm thick porcine tissue by using surface-enhanced spatially offset Raman spectroscopy. *Angew. Chem.* **124**, 8637-8639, (2012).

193   Sharma, B., Ma, K., Glucksberg, M. R. & Van Duyne, R. P. Seeing through bone with surface-enhanced spatially offset Raman spectroscopy. *J. Am. Chem. Soc.* **135**, 17290-17293, (2013).

194   Saar, B. G. *et al.* Label-free, Real-time monitoring of biomass processing with stimulated Raman scattering microscopy. *Angew. Chem. Int.* **49**, 5476-5479, (2010).

195   Littlejohn, G. R. *et al.* In vivo chemical and structural analysis of plant cuticular waxes using stimulated Raman scattering (SRS) microscopy. *Plant Physiol.* **168**, 18-28, (2015).

196   Kneipp, K. *et al.* Single molecule detection using surface-enhanced Raman scattering (SERS). *Phys. Rev. Lett.* **78**, 1667-1670, (1997).

197   Nie, S. & Emory, S. R. Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science (New York, N.Y.)* **275**, 1102-1106, (1997).

198   Qian, X. *et al.* In vivo tumor targeting and spectroscopic detection with surface-enhanced Raman nanoparticle tags. *Nat. Biotechnol.* **26**, 83-90, (2008).

199   Kneipp, K., Kneipp, H. & Kneipp, J. Surface-enhanced Raman scattering in local optical fields of silver and gold nanoaggregates from single-molecule Raman spectroscopy to ultrasensitive probing in live cells. *Accounts Chem. Res.* **39**, 443-450, (2006).

200   Cîntă Pînzaru, S., Pavel, I., Leopold, N. & Kiefer, W. Identification and characterization of pharmaceuticals using Raman and surface-enhanced Raman scattering. *J. Raman Spectrosc.* **35**, 338-346, (2004).

201   Huang, X., El-Sayed, I. H., Qian, W. & El-Sayed, M. A. Cancer cells assemble and align gold nanorods conjugated to antibodies to produce highly enhanced, sharp, and polarized surface Raman spectra: A potential cancer diagnostic marker. *Nano Lett.* **7**, 1591-1597, (2007).

202   Feng, S. *et al.* Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. *Biosens. Bioelectron.* **25**, 2414-2419, (2010).

203   Premasiri, W. *et al.* Characterization of the surface enhanced Raman scattering (SERS) of bacteria. *J. Phys. Chem. B* **109**, 312-320, (2005).

204   Zeiri, L. SERS of plant material. *J. Raman Spectrosc.* **38**, 950-955, (2007).

205   McNay, G., Eustace, D., Smith, W. E., Faulds, K. & Graham, D. Surface-enhanced Raman scattering (SERS) and surface-enhanced resonance Raman scattering (SERRS): a review of applications. *Appl. Spectrosc.* **65**, 825-837, (2011).

206   Graham, D. *et al.* Selective detection of deoxyribonucleic acid at ultralow concentrations by SERRS. *Anal. Chem.* **69**, 4703-4707, (1997).

207   Murgida, D. H. & Hildebrandt, P. Electron-transfer processes of cytochrome c at interfaces. New insights by surface-enhanced resonance Raman spectroscopy. *Accounts Chem. Res.* **37**, 854-861, (2004).

208   Murgida, D. H. & Hildebrandt, P. Disentangling interfacial redox processes of proteins by SERR spectroscopy. *Chem. Soc. Rev.* **37**, 937-945, (2008).

209   Böhme, R. *et al.* Biochemical imaging below the diffraction limit–probing cellular membrane related structures by tip-enhanced Raman spectroscopy (TERS). *J. Biophotonics* **3**, 455-461, (2010).

210   Stone, N. & Matousek, P. Advanced transmission Raman spectroscopy: a promising tool for breast disease diagnosis. *Cancer Res.* **68**, 4424-4430, (2008).

211   Johansson, J., Sparén, A., Svensson, O., Folestad, S. & Claybourn, M. Quantitative transmission Raman spectroscopy of pharmaceutical tablets and capsules. *Appl. Spectrosc.* **61**, 1211-1218, (2007).

212   Buckley, K. & Matousek, P. Recent advances in the application of transmission Raman spectroscopy to pharmaceutical analysis. *J. Pharm. Biomed. Anal.* **55**, 645-652, (2011).