A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology

To cite this article: Michael Jermyn et al 2016 Phys. Med. Biol. 61 R370

View the article online for updates and enhancements.

Related content
- Biophotonics: Vibrational Spectroscopic Diagnostics: Vibrational biomedical spectroscopy of tissue
  M J Baker, C S Hughes and K A Hollywood
- Prospects for in vivo Raman spectroscopy
  E B Hanlon, R Manoharan, T-W Koo et al.
- Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy
  Kenny Kong, Fazliyana Zaabar, Emad Rakha et al.

Recent citations
- Empirical Factors Affecting the Quality of Non-Negative Matrix Factorization of Mammalian Cell Raman Spectra
  H. Georg Schulze et al
- Raman spectroscopy as a novel method in placental research: Recognizing the pattern of placental hypoxia
  Natalia Schlabritz-Loutsevich et al
- Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics
  Inês P. Santos et al

Versa HD™
Powered by high definition dynamic radiosurgery.

Click here to learn more Versa HD
Topical Review

A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology

Michael Jermyn\textsuperscript{1,2,3}, Joannie Desroches\textsuperscript{2}, Kelly Aubertin\textsuperscript{3}, Karl St-Arnaud\textsuperscript{2}, Wendy-Julie Madore\textsuperscript{2}, Etienne De Montigny\textsuperscript{2}, Marie-Christine Guiot\textsuperscript{4}, Dominique Trudel\textsuperscript{5,6,7}, Brian C Wilson\textsuperscript{8}, Kevin Petrecca\textsuperscript{1} and Frederic Leblond\textsuperscript{2,3,9}

\textsuperscript{1} Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, 3801 University St., Montreal, QC H3A 2B4, Canada
\textsuperscript{2} Department of Engineering Physics, Polytechnique Montreal, CP 6079, Succ. Centre-Ville, Montreal, QC H3C 3A7, Canada
\textsuperscript{3} Imaging and Engineering, University of Montreal Hospital Research Center (CR-CHUM), 900 Rue Saint-Denis, Montréal, QC H2X 0A9, Canada
\textsuperscript{4} Department of Pathology, Division of Neuropathology, McGill University, 3801 University St., Montreal, QC H3A 2B4, Canada
\textsuperscript{5} Department of Pathology, Centre Hospitalier Universitaire de Montréal, Montreal, QC H2X 3J4, Canada
\textsuperscript{6} University of Montreal Hospital Research Center (CRCHUM)/Montreal Cancer Institute, 900 Rue Saint-Denis, Montréal, QC H2X 0A9, Canada
\textsuperscript{7} Department of Pathology and Cellular Biology, University of Montreal, Montreal, QC H3T 1J4, Canada
\textsuperscript{8} Princess Margaret Cancer Centre, 101 College St., Toronto, ON M5G 1L7, Canada

E-mail: frederic.leblond@polymtl.ca

Received 10 August 2015, revised 22 June 2016
Accepted for publication 29 September 2016
Published 2 November 2016

Abstract

There is an urgent need for improved techniques for disease detection. Optical spectroscopy and imaging technologies have potential for non- or minimally-invasive use in a wide range of clinical applications. The focus here, in vivo Raman spectroscopy (RS), measures inelastic light scattering based on interaction with the vibrational and rotational modes of common molecular bonds in cells and tissue. The Raman ‘signature’ can be used to assess physiological status and can also be altered by disease. This information can supplement existing diagnostic (e.g. radiological imaging) techniques for disease screening and diagnosis, in interventional guidance for identifying disease margins, and in monitoring treatment responses. Using
fiberoptic-based light delivery and collection, RS is most easily performed on accessible tissue surfaces, either on the skin, in hollow organs or intra-operatively. The strength of RS lies in the high biochemical information content of the spectra, that characteristically show an array of very narrow peaks associated with specific chemical bonds. This results in high sensitivity and specificity, for example to distinguish malignant or premalignant from normal tissues. A critical issue is that the Raman signal is often very weak, limiting clinical use to point-by-point measurements. However, non-linear techniques using pulsed-laser sources have been developed to enable in vivo Raman imaging. Changes in Raman spectra with disease are often subtle and spectrally distributed, requiring full spectral scanning, together with the use of tissue classification algorithms that must be trained on large numbers of independent measurements. Recent advances in instrumentation and spectral analysis have substantially improved the clinical feasibility of RS, so that it is now being investigated with increased success in a wide range of cancer types and locations, as well as for non-oncological conditions. This review covers recent advances and continuing challenges, with emphasis on clinical translation.

Keywords: Raman spectroscopy, surgical guidance, tissue diagnosis, optical guidance, spectral imaging

(Some figures may appear in colour only in the online journal)

Introduction

Raman spectroscopy (RS) is a label-free optical technique with the capability of non- or minimally-invasively determining molecular information from biological tissue. Inelastic scattering of light was first observed by Indian Nobel Laureate C V Raman in 1928 (Raman and Krishnan 1928). However, the phenomenon was not reported for biomedical applications until 1970 (Lord and Yu 1970). Subsequent improvements in light sources and signal detection have made possible the emergence of RS techniques across a wide range of biological applications.

The vast majority of light scattering in tissue is elastic, i.e. there is no energy exchange between the photons and the molecules (Rayleigh scattering). However, in tissue typically only one in several million photons undergoes inelastic scattering through energy exchange with the vibrational or rotational modes of molecular bonds (Ferraro et al 2002). This low probability makes detection of Raman-scattered light challenging, especially on the large background of elastic scattering as well as fluorescence and phosphorescence from the tissue (Zhao et al 2007, Dochow et al 2012). The Raman cross section varies as $1/\lambda^4$, where $\lambda$ is the wavelength of the incident light. Thus, UV or short-wavelength visible light gives the largest Raman signal from tissue. However, this is more than offset by the high fluorescence background, which falls off markedly in the near-infrared range. Compared with fluorescence from biomolecules, which is characterized by a few, rather broad features (tens of nm wide), Raman spectra from tissue typically have multiple sharp peaks (~few nm wide), allowing the broad fluorescence background to be subtracted. The spectral shift between the incident light and the RS light is also small, making spectral separation from the dominant elastic-scattering light difficult. However, technology advances over the past 20 years have led to increased signal and contrast, enabling shorter and, hence, clinically practical acquisition times. These have included stable diode laser sources, high-throughput and high-resolution spectrographs
and sensitive array detectors. The clinical practicality has also been markedly enhanced by the use of fiberoptic light delivery and collection, as well as small and high-efficiency spectral filters to remove background light. In addition, powerful spectral analysis and classification algorithms have been developed and implemented on portable computers.

In oncology, RS has been shown to be highly sensitive to the altered 'molecular signatures' of many different cancers. Despite advances in early detection and treatment, cancer remains one of the leading causes of death, with ~14 million new cancer cases and ~8 million cancer-related deaths globally in 2012 (Ferlay et al 2015). Moreover, the burden of cancer is expected to increase in the coming decades including in the developing world where access to affordable, point-of-care medical technology is problematic. Current diagnostic methods for cancer typically involve the use of histopathology of biopsied or resected tissue (or cytology of blood in the case of hematological malignancies) and radiological imaging modalities such as magnetic resonance imaging (MRI), ultrasound (US), and computed tomography (CT), and gamma-ray imaging and positron emission tomography. However these methods are expensive, may be time consuming, may not be optimal for intraoperative use and often fail to fully characterize the extent of cancer or detect smaller tumors. Hence, there remains an urgent need for improved cancer detection, characterization and localization technologies. Here we are reviewing in detail progress in RS for cancer applications. RS has also been investigated for a number of other clinical applications, including in Alzheimer's disease (Ryzhikova et al. 2015), corneal hydration (Boncheva et al. 2009), nephropathic cystinosis (Cinotti et al. 2013) and cardiovascular diseases (Motz et al. 2006, Cloyd et al. 2012).

The minimally-invasive nature of RS and the use of endogenous (i.e. label-free) contrast, in particular, make it a very attractive and safe approach for in vivo use. However, as indicated above, significant challenges remain for routine clinical application of RS, in large part because of the rarity of the inelastic scattering process in tissue. The Raman signal can be markedly increased through the use of metal nanoparticles (surface enhanced Raman scattering, SERS) or by exploiting non-linear scattering interactions with ultrashort-pulsed laser sources (coherent anti-Stokes Raman scattering, CARS; stimulated Raman scattering, SRS), and the status of these approaches for clinical applications will be discussed.

Tissues also are comprised of a very large number of molecular species that may have similar Raman signatures. In particular, RS does not usually identify individual molecular species. Rather, the spectral features are due to molecular bonds (e.g. CH, CH₃, etc) that are common to many different biomolecules. Thus, molecular changes with disease usually alter multiple features across the spectral range, so that sophisticated spectral analysis techniques are required for tissue classification, e.g. to distinguish between 'disease' and 'normal' or between different stages of disease progression. These algorithms then require training, usually by comparison with the gold standard of histopathology. This is, however, not without its pitfalls, and may actually limit the apparent accuracy that can be achieved with RS, as will be discussed.

Clinicians must be provided with unambiguous data to help make informed decisions. Furthermore, disease progression plays an important role in the molecular diagnostic information of RS. Moving Raman techniques toward clinical use will therefore require multidisciplinary interactions among engineers, clinicians, and other specialists. In this review we will discuss the challenges and recent advances for the clinical use of RS for disease detection and also provide some indication for its synergistic use with technologies exploring other optical contrast mechanisms. Although Raman spectroscopy and imaging are increasingly used as research tools, for example, in the analysis of cellular and animal models of disease, and these studies are often used to inform clinical trial design and/or interpretation, the focus of this review will be on in vivo RS as it can be applied for patients in the clinical environment.
Raman techniques for disease detection

The following sections will present a brief overview of each technique as it has been developed and applied to date to address biomedical, and particularly clinical, problems. Further reading can be found in previous reviews (Krafft et al 2012b, Antonio and Schultz 2014, Rae et al 2014, Austin et al 2016). Table 1 provides a detailed comparison of the various techniques. Note that while some of the highlighted features are intrinsic to the techniques, some will depend on advancements in technology. The approaches are broadly divided into: ‘spontaneous’, corresponding to the classical linear interactions; ‘coherent’, corresponding to enhanced, non-linear interactions; and ‘non label-free’, referring to the enhanced signal from biomolecules close to metal (nano-structured) surfaces.

Spontaneous Raman spectroscopy (SpRS)

Under beam illumination at a given frequency, a molecule scatters most of the light elastically at the same frequency. The molecule, initially in the ground state, is promoted to a virtual state and reemits a photon of the same frequency to return to its initial state (Rayleigh scattering). Inelastic Raman scattering produces a shift in frequency proportional to the difference of energy between the vibrational states of the molecule (Suhalim et al 2012). The molecule promoted to a virtual state can move from a vibrational state to the ground state, or move from the ground state to a vibrational state. These are referred to, respectively, as the anti-Stokes transition where the scattered light is at higher frequency (lower energy) and the Stokes transition where the scattered light at lower frequency (higher energy) than the incident photon. The latter is typically much stronger and so more commonly used. Raman analysis of a complex molecule will yield a spectrum of spectral shifts where the peaks are associated with vibrational and rotational modes in the molecule. For tissue, the peak positions in a spectrum are characteristic of the molecular composition of tissue and the amplitude of each peak depends on the relative concentration of each molecule.

In ‘conventional’ spontaneous RS (hereafter referred to as simply RS), the Stokes signal is detected under continuous wave (CW) illumination, typically from a diode laser with high spectral stability. As mentioned above, the signal intensity increases for shorter wavelengths but the tissue intrinsic fluorescence also increases markedly, while the penetration depth of light in the tissue decreases due to absorption of intrinsic chromophores, particularly hemoglobin and melanin, so that the RS signal-to-background is reduced. Moreover, the absorption spectrum, \( \mu_\lambda(\lambda) \), is strongly wavelength- and tissue-dependent in the visible range, which confounds comparisons of the RS signal intensity. Hence, the optimal optical windows for RS in tissue are between 700 and 900 nm and around 1064 nm. In the literature, 785, 830 and 1064 nm are primarily used (Kelly et al 2012). The choice of wavelength typically involves a trade-off between signal intensity and tissue autofluorescence and depend on the particular tissues involved and the application. SpRS techniques generally use a spectrometer for detection. Since Rayleigh scattering is much larger than Raman scattering, filtering is necessary before detection (Pence and Mahadevan-Jansen 2016).

For in vivo clinical RS, it is most convenient to deliver the light to the tissue and collect the RS light using a specifically-designed fiber probe, and several different probe designs for point-based measurements in contact with tissue have been reported (Mahadevan-Jansen et al 1998, Haka et al 2009, Beljebbar et al 2010, Lui et al 2012, Jermyn et al 2015). The probe typically has one central illumination fiber surrounded by a number of detections fibers, with a lens at the tip to maximize collection of Raman signal. One of the biggest challenges for probe design is to minimize the detection of tissue fluorescence and the fluorescence and
<table>
<thead>
<tr>
<th>Optics Technique</th>
<th>Light source (imaging technique)</th>
<th>Acquisition time</th>
<th>Spatial resolution</th>
<th>Signal/background</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpRS</td>
<td>Continuous laser (probe or microscope)</td>
<td>&gt;200 ms</td>
<td>&gt;0.2 mm²</td>
<td>Signal intensity $\propto I_p N_{SERS}$</td>
<td>- High background</td>
</tr>
<tr>
<td></td>
<td>Frequency $\omega = \omega_p$ (500–1100 nm)</td>
<td></td>
<td></td>
<td>Signal frequency $\omega_S = \omega_p - \delta$</td>
<td>- Low axial resolution</td>
</tr>
<tr>
<td></td>
<td>Average power $&lt;150$ mW</td>
<td></td>
<td></td>
<td>Background Intrinsic tissue autofluorescence</td>
<td>- Inexpensive</td>
</tr>
<tr>
<td></td>
<td>Point-based full spectrum</td>
<td></td>
<td></td>
<td></td>
<td>- Relatively simple</td>
</tr>
<tr>
<td></td>
<td>&gt;5 s</td>
<td></td>
<td></td>
<td></td>
<td>- Point, line or array imaging</td>
</tr>
<tr>
<td></td>
<td>Point-based full spectrum</td>
<td>&gt;2 µm</td>
<td></td>
<td></td>
<td>- High background</td>
</tr>
<tr>
<td></td>
<td>&gt;5 s</td>
<td></td>
<td></td>
<td></td>
<td>- High axial resolution</td>
</tr>
<tr>
<td></td>
<td>Point-based full spectrum</td>
<td>&gt;5 mm²</td>
<td></td>
<td></td>
<td>- Deep tissue analysis (&lt;500 µm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Long acquisition time</td>
</tr>
<tr>
<td>SORS (spatially offset RS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional RS</td>
<td>Non label-free RS</td>
<td>Same as SpRS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SERS (surface enhanced RS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;1 s</td>
<td></td>
<td></td>
<td></td>
<td>- Low background</td>
</tr>
<tr>
<td></td>
<td>Point-based full spectrum</td>
<td>&gt;1 µm</td>
<td></td>
<td></td>
<td>- Requires biomarkers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Non-uniform signal intensity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Requires surface coverage</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Long acquisition time</td>
</tr>
<tr>
<td>Light source (imaging technique)</td>
<td>Acquisition time</td>
<td>Spatial resolution</td>
<td>Signal/background</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Coherent RS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CARS (coherent anti-Stokes RS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 pulsed lasers or 1 pulsed laser with optical parametric oscillator (OPO)</td>
<td>$&gt;0.16 \mu s/cm^{-1}$ Pixel dwell time</td>
<td>Lateral $&gt;0.3 \mu m$ Axial $&gt;1.5 \mu m$ Image: $800 \times 600 \mu m$</td>
<td>Signal intensity $\propto</td>
<td>\chi^{(3)}</td>
<td>I_p^2N^2$</td>
</tr>
<tr>
<td>SRS (stimulated RS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average power $\leq 150 \text{mW}$ Pulse $\sim 5\text{ps}$ Repetition rate 1–100 MHz</td>
<td>$&gt;37\text{ms/frame}$ Image: $512 \times 512$</td>
<td>Signal intensity $\propto \text{Im}(</td>
<td>\chi^{(3)}</td>
<td>)I_pN$ Signal frequency $(\omega_{SRL} \text{and } \omega_{SRG})$ $\omega_{SRL} = \omega_i + \delta = \omega_p$ $\omega_{SRG} = \omega_p - \delta = \omega_k$ Background No intrinsic tissue autofluorescence background</td>
</tr>
</tbody>
</table>

Note. The parameter thresholds reported for each technique represent the best values achieved in the literature for biological tissue to our knowledge (Shafer-Peltier et al. 2002, Evans et al. 2005, Stone et al. 2007, Saar et al. 2010, Keller et al. 2011, Bergner et al. 2012, Dinish et al. 2014, Jermyn et al. 2015). Values for light source power are for point illumination of tissue. Definitions for terms used: pump frequency ($\omega_p$), Stokes frequency ($\omega_s$), anti-Stokes frequency ($\omega_{AS}$), stimulated Raman loss frequency ($\omega_{SRL}$), stimulated Raman gain frequency ($\omega_{SRG}$), pump laser intensity ($I_p$), Stoke laser intensity ($I_s$), Number of molecules ($N$), number of molecules for SERS ($N'$), Raman scattering cross section ($\sigma_{\text{Raman}}$), SERS cross section ($\sigma_{\text{SERS}}$), enhancement factor at Stokes frequency ($A_s$), enhancement factor at pump frequency ($A_p$), and third-order susceptibility ($\chi^{(3)}$).
Raman-scattered light generated in the probe itself, such as the silica Raman of glass-fiber cores and the fluorescence and Raman from the fiber buffer/cladding. These background signals can be generated both in the light-delivery fiber and, unless the light that is elastically-scattered by the tissue is blocked, also in the collection fibers. Incorporating appropriate spectral filters into the probe is particularly challenging (and expensive) if the probe diameter needs to be very small (<~1–2 mm) so that it can pass down the instrument channel of an endoscope. Such SpRS probes typically sample a tissue volume of <1 mm² laterally by a few hundred microns axially (Jermyn et al 2015). There are a number of probe designs optimized for different clinical applications, including contact probes, endoscopic probes, and needle-based probes (Stevens et al 2016). It is also possible to use Raman microscopy for wide-field imaging, however acquisition time is often prohibitively long for in vivo applications of SpRS.

Confocal Raman spectroscopy

SpRS can also be implemented in confocal configuration to provide optical depth sectioning, by spatially filtering the collected RS light with a pinhole or an optical fiber to block out-of-focus signal. This also improves the lateral and axial resolutions to as low as ~2 µm (Shafer-Peltier et al 2002). However, the acquisition time per point is usually >5 s and increases with focal depth in the tissue. Confocal Raman probes can achieve acquisition times as low as 1 s but only with significant loss in lateral resolution (Wood et al 2014). Raster-scanning confocal RS yields 2D images with a field-of-view typically ranging between 0.01 and 1 mm². Confocal Raman probes have to date been used mainly for ex vivo and in vitro studies (Day et al 2009, Wang et al 2013b, Wood et al 2014).

Spatially offset Raman spectroscopy (SORS)

SORS is similar to SpRS, but collects Raman signal from deeper regions in tissue by spatially offsetting the detection and excitation fibers. The detected photons will then have been elastically scattered multiple times and traverse some distance from the illumination source. Collecting the Raman signal at different offsets effectively samples different layers in the tissue. SORS typically uses a probe with an illumination fiber surrounded by detections fibers offset by 1–5 mm (Stone et al 2007, Macleod et al 2008, Keller et al 2009, 2011), but an offset as high as 16 mm has been used to perform Raman tomographic imaging in bone (Schulmerich et al 2008). Since SORS collects signal at depth, the signal-to-noise ratio is usually reduced, so that integration times of 10–30 s are required (Keller et al 2009, 2011). A breast cancer study reported detection up to depths of 2 mm, which has the potential to improve tumor margin assessment during surgery (Keller et al 2011). Penetration depths up to 10 mm have been reported for detecting calcifications in chicken breast as a model tissue (Stone et al 2007).

Coherent Raman spectroscopy (coherent RS)

Coherent RS uses two light fields (referred to as the pump and Stokes beam) with frequencies ωp and ωs respectively, so that the difference corresponds to a vibrational mode frequency, Ω = ωp − ωs for the molecular bond of interest. The coherent addition of the Raman signal from different molecules improves the signal compared to SpRS, typically by up to ~10⁵ fold (Kong et al 2015). A number of techniques can be used to sweep the frequency (Alfonso-García et al 2014). The improved signal from coherent RS may be especially important for fast imaging, e.g. at video rate with 30 frames s⁻¹ and pixel dwell times as low as 100 ns.
Coherent RS techniques include stimulated Raman spectroscopy (SRS) and coherent anti-Stokes Raman spectroscopy (CARS). Both SRS and CARS can be performed in highly fluorescent media, which is usually a significant limiting factor for Raman imaging in tissue.

CARS uses a wave-mixing process to generate the anti-Stokes Raman frequency, \( \omega_{\text{AS}} = 2\omega_p - \omega_s \). When the frequency \( \omega_p - \omega_s \) matches one of the molecular vibration states, the mode is coherently driven and signal intensity increases at the anti-Stokes frequency. CARS has been used to detect molecular vibrations associated with lipids \( \text{(Langohr et al. 2007, Lim et al. 2011, Meyer et al. 2011)} \) and has also been performed \( \text{(Evans et al. 2005)} \), while probe-based techniques have been proposed \( \text{(Légaré et al. 2006, Saar et al. 2011)} \). However, there are challenges in probe designs for CARS \( \text{(Balu et al. 2010)} \), due to pulse broadening within the delivery optical fiber as well as other non-linear optical effects resulting from the very high power densities used. Strong anti-Stokes light from the fiber must be filtered before signal acquisition. A further limitation of CARS is the presence of a significant non-resonant background that adds coherently with the resonant signal \( \text{(Zheng et al. 2015)} \). This is particularly an issue in tissue since water has a large non-resonant signal. Moreover, CARS spectra can differ from spontaneous Raman spectra due to the non-resonant background, and because the CARS signal varies non-linearly with molecular concentration \( \text{(Alfonso-García et al. 2014)} \).

As in CARS, SRS uses two synchronized lasers at frequencies \( \omega_p \) (pump) and \( \omega_s \) (Stokes). For SRS, the pump laser stimulates the molecule to a virtual state. Then, when the Stokes laser matches a vibrational state of the molecule, it can stimulate photon emission at the same frequency, \( \omega_s \). The detection setup for SRS is more complex than with CARS and, since the variation in intensity is very small, a high modulation rate is needed to overcome low-frequency fluctuations of the laser intensity. Nevertheless, SRS has shown great potential for video rate imaging, with up to 30 frames s\(^{-1}\) for \( 512 \times 512 \) pixel images \( \text{(Saar et al. 2010)} \). This system required a pixel dwell time of \( 2 \mu s \) for discriminating tumor tissue \textit{in vivo} in mice \( \text{(Ji et al. 2013)} \), so that it may not be possible to operate at maximum frame rate in practice. Multiplexed video rate imaging has also been performed \( \text{(Freudiger et al. 2011, Fu et al. 2012)} \). A recent study found that spatial-frequency multiplexing with single photodiode detection for SRS offered improved photon collection efficiency \textit{in vivo} for mouse skin and \textit{in situ} for human breast cancer, with <60 \( \mu s \) acquisition for \( 16 \) frequency components \( \text{(figure 1)} \) \( \text{(Liao et al. 2015)} \). This approach should accelerate clinical translation of SRS imaging, particularly for rapid \textit{in vivo} assessment of cancerous tissue. One of the main advantages of SRS over CARS is that stimulated emission can only occur if \( \omega_p - \omega_s \) matches a vibrational state of the molecule. Consequently, there is no non-resonant background and SRS spectra are more easily compared with SpRS spectra. Although SRS is free from intrinsic autofluorescence and non-resonant background, some non-linear optical effects can occur quasi-instantaneously, thereby producing background SRS signal \( \text{(Berto et al. 2014)} \).

Among the challenges for \textit{in vivo} translation of coherent RS techniques are to achieve efficient light delivery and collection using optical fibers, and to miniaturize the optical components for endoscopic or contact-probe use \( \text{(Latka et al. 2013)} \).

**Surface enhanced Raman spectroscopy (SERS)**

Low signal intensity is clearly a major limitation of RS. An alternative to coherent amplification of the signal is to use effects occurring near a metal surface, including metal nanoparticles, to produce electromagnetic and chemical amplification \( \text{(Schlücker 2014)} \). The former is usually the dominant factor; under an incident electromagnetic field metal nanoparticles generate...
Topical Review

a localized surface plasmon resonance that enhances the illumination at the pump frequency and the Raman signal at the Stokes frequency. The enhancement falls off very rapidly with distance, so that the molecules must be within tens of nanometers of the surface (Dieringer et al 2006, Shanthil et al 2012, Schlücker 2014). The chemical enhancement is much weaker and is due to wavelength-specific resonance with the charge transfer between the nanoparticles and the molecule. There have been two main approaches to SERS in biomedical applications. The first is to coat an optical fiber probe with metal nanoparticles that is then placed on or in the biological sample and amplifies the intrinsic Raman signature (Schlücker 2014). The second is to administer metal nanoparticles that have been coated with reporter molecules of known Raman spectra. The nanoparticles may then be targeted to specific biomarkers, for example, using antibodies or peptides, acting as nanotags. Since Raman spectral peaks are so narrow and molecule specific, multiplexed detection can be achieved by using a grouping of several different nanoparticles, each with a different combination of reporter and targeting molecules. Simultaneous multiplexed imaging of up to 10 nanotags have been reported (Zavaleta et al 2009, Schütz et al 2011, Dinish et al 2014), while McVeigh et al have demonstrated that spectral decomposition of SERS multiplexed images can be quantitative (McVeigh et al 2013), i.e. represent the true relative concentration of the biomarkers: in part this is due to the high signal strength that reduces the effect of tissue autofluorescence background (Feng et al 2010). Limitations of SERS for clinical applications are that it depends on knowledge of sensitive disease biomarkers and the availability of corresponding targeting moieties, as well as potential toxicity and the need for regulatory approval of the contrast agent.

Figure 1. (a) Imaging of human breast cancer tissue with single photodiode detection of SRS. (b) Spectra for fat, fibrosis, and cell nuclei. (c) SRS images at 2908 cm⁻¹ for corresponding locations in the tissue. (d) Calculated concentration maps of fat (yellow), fibrosis (green), and cell nuclei (pink). (e) Corresponding histological images. From Liao et al (2015). Reprinted with permission from AAAS. This work is licensed under CC BY-NC (CC BY-NC 4.0).
Diagnostic applications

Raman spectroscopy can be used for a number of clinical applications, including surgical guidance, screening and histopathology. It is increasingly investigated for tissue characterization, enabled by recent technological and analytical advances. For example, a number of companies have already developed in vivo systems for clinical use of RS, such as: Verisante (Canada) for skin cancer, Endofotonics (Singapore) for endoscopic cancer detection and ODS Medical (Canada) for brain cancer. RS has also seen increased use for disease detection via the measurement of biofluids such as blood, saliva, and urine, with several companies developing toolkits and microscopes for this purpose, e.g. RiverD (The Netherlands). While there have been dozens of single center studies using Raman technology in human applications, there have thus far been a very limited number of prospective studies, no blinded clinical studies, and no multicenter studies conducted to evaluate clinical impact. Moreover, no companies have obtained regulatory approval (e.g. FDA in the USA, CE Mark in Europe) for the use of Raman-based instruments in clinical practice, with the exception of Verisante for skin lesion diagnosis (Canada and Europe). Here, we present recent research for diagnostic applications of RS with an emphasis on work that has already shown surgical guidance applications or has significant potential to be translated into the operating theater. Unless otherwise stated, conventional spontaneous RS has been the method of choice in these studies.

Skin

One of the most advanced applications of RS is for the detection of skin cancer, facilitated by ease of access. This includes melanomas, basal cell carcinomas (BCC) and squamous cell carcinomas (SCC). Clinical diagnosis of skin cancer is conventionally based mainly on visual inspection followed by biopsy and histological analysis of suspicious lesions. Macroscopic evaluation of lesions is highly subjective and strongly depends on the clinician’s experience, while biopsies are subject to sampling errors, leading to a large number of false negatives and unnecessary biopsies. The use of RS for skin cancer screening could help to reduce these diagnostic errors, in addition to identifying cancer cells that invade normal tissue at the periphery of a lesion, i.e. better defining the true tumor margin to guide excision.

While many groups have used RS on in vitro and ex vivo samples, recent in vivo studies also show promising results.

Zhao et al measured in vivo Raman spectra on 289 patients presenting 9 different lesion types including BCC and SCC and achieved 91% sensitivity and 75% specificity for differentiating cancers from benign lesions, and 97% sensitivity and 78% specificity for distinguishing malignant melanoma from pigmented benign lesions (Zhao et al 2008). It was thought that skin pigmentation might be a confounding factor in the use of RS for skin cancer diagnosis. Following two ex vivo studies, the Wulf group published an in vivo study in 55 patients showing that skin pigmentation resulted in a higher Raman background due to intrinsic fluorescence, but that this can be subtracted using standard algorithms (Gniadecka et al 2004, Sigurdsson et al 2004, Philipson et al 2013). They also demonstrated significant differences ($p < 0.05$) in the Raman bands corresponding to water, proteins and the protein/lipid ratio. Schleusener et al were unsuccessful in differentiating cancerous form benign skin lesions but were able to discriminate malignant melanoma from pigmented nevi in vivo with 87% sensitivity and 94% specificity (Schleusener et al 2015).
Brain Gliomas represent 80% of adult malignant brain tumors and aggressively invade into normal brain (Goodenberger and Jenkins 2012). The standard for surgical resection is visual inspection through a neurosurgical microscope along with navigational guidance from magnetic resonance imaging (MRI). However, the full extent of tumor infiltration is often not detected, leading to residual tumor and recurrence. Moreover, the removal of normal tissue can cause permanent neurological deficits. RS has the potential to be useful for neurosurgical guidance by increasing the ability to detect residual tumor 
\textit{in vivo} and extend safe resection. Many groups have investigated the spectral differences between normal and cancerous tissue using \textit{ex vivo} human tissue or rodent glioma models and have revealed important biochemical differences (Mizuno \textit{et al} 1994, 1992, Koljenović \textit{et al} 2002, 2007, Krafft \textit{et al} 2006, 2012a, Beljebbar \textit{et al} 2010, Kirsch \textit{et al} 2010, Kalkanis \textit{et al} 2014). For example, Beljebbar \textit{et al} implanted a glioblastoma model in rats and used a portable Raman spectrometer coupled to a microprobe to show that the molecular differences between normal brain tissue, tumor invasion and frank tumor tissue gave a classification accuracy of 100% (Beljebbar \textit{et al} 2010). They were also able to track the evolution of the spectra during the tumor development. SRS has been used to differentiate tumor from non-neoplastic tissue in infiltrative human glioblastoma xenografts in mouse brain, with the contrast based on the signal intensities at 2930 and 2845 cm$^{-1}$ (figure 2) (Ji \textit{et al} 2013). During \textit{in vivo} resection, SRS microscopy was able to reveal extensive tumor infiltration not detectable under normal bright-field microscopy.

\textbf{Figure 2.} (a) \textit{In vivo} bright-field and SRS microscopy of human glioblastoma xenografts in mouse brain, showing tumor-specific differences in SRS not visible in the bright-field image. The dashed line indicates the boundary between tumor and normal brain. The SRS images are false colored, with lipids colored green and proteins colored blue based on the signals at 2845 and 2930 cm$^{-1}$, respectively. (b) High-magnification view of the tumor. (c) High-magnification view of the tumor-brain interface. (d) High-magnification view of normal brain. From Ji \textit{et al} (2013). Reprinted with permission from AAAS.
et al used RS for in vivo imaging of cerebral brain metastases in mice, with acquisition times of 4 s per spectrum, or several minutes per image (figure 3) (Kirsch et al 2010). A system using a hand-held RS probe was developed and optimized for intraoperative use during brain tumor resection with 0.2 s acquisition time (Desroches et al 2015, Jermyn et al 2015). The system was used in vivo on 17 patients with grade 2–4 gliomas for a total of 161 spectra, yielding 93% sensitivity and 91% specificity for distinguishing normal brain from dense cancer and low-density tumor infiltration.

Breast

RS can be used during breast-conserving surgery (lumpectomy) which is usually followed by radiotherapy. The surgical outcome of surgery depends strongly on the presence of positive margins after tumor removal. Local recurrence, requiring second surgery carries risk, cost and patient stress, so that intraoperative detection of positive margins represents a significant unmet clinical need. Standard margin assessment using frozen sections is time consuming, labor intensive and is not sufficiently accurate. A number of techniques are under development, including optical methods such as optical coherence tomography (South et al 2014) and non-optical such as radiofrequency spectroscopy (Thill 2013). RS may address the need for rapid and robust intraoperative breast cancer margin assessment.
There are numerous studies on the application of RS to breast cancer diagnosis, mostly using *ex vivo* human tissue samples or animal models. Bhattacharjee *et al* used transcutaneous *in vivo* RS in a mouse and achieved a classification accuracy of 99% with a 15 s integration time (Bhattacharjee *et al* 2013). The first *in vivo* RS study, using <1 s integration time, in patients was reported by Haka *et al* in 9 patients undergoing partial mastectomy and claimed 100% accuracy in classifying tissue as tumor versus non-tumor. However, the statistical significance of this result was limited (Haka *et al* 2006). Three years later, the same group published a prospective analysis on *ex vivo* samples showing 83% sensitivity and 93% specificity (Haka *et al* 2009).

**Gynecology**

Ovarian cancer is one of the leading causes of death in women in developed countries. It is difficult to diagnose early, since no effective screening tests are available and symptoms are vague. In addition, tumor can spread easily throughout the uterine corpus, fallopian tube, ovaries, peritoneal cavity and/or abdominal cavity. Multiple ovarian cancer biomarkers have been studied as potential screening tools, but to date none have adequate sensitivity and specificity. This motivates investigation of optical detection techniques.

Krishna *et al* identified RS peaks specific to malignant ovarian tissue and established that Raman spectral features of ovarian tissue have less fixation artifacts when fixed in formalin than when embedded in paraffin (Krishna *et al* 2005). A subsequent study looked at healthy, benign and malignant formalin-fixed ovarian tissue from surgical resections (Krishna *et al* 2007) and found that benign and healthy tissue spectra are similar, and are clearly differentiated from the malignant tissue spectra. Maheedhar *et al* achieved 100% sensitivity and specificity for detecting ovarian cancer in 72 spectra from 15 patients on freshly excised ovaries, with acquisition times of several minutes per spectrum (Maheedhar *et al* 2008). Jokerst *et al* explored the potential of SERS for surgical resection of ovarian tumors in a murine model. By injecting gold nanorods 3 h prior to surgery, they were able to perform resection guided by SERS *in vivo* by discriminating tumor margins (Jokerst *et al* 2012). Unfortunately, high liver uptake of the nanoparticles introduced additional non-specific Raman background. Boca-Farcau *et al* reported a more specific marker: silver nanotriangles that were both SERS labeled and folic-acid conjugated (Boca-Farcau *et al* 2014). Since folate receptors are known to be overexpressed in most ovarian epithelial cancers, these nanoparticles have potential to target ovarian cancer in the surgical field with higher specificity. A study in human cell lines showed good intracellular internalization after 24 h incubation. RS has also been used to analyze uterine tissue for endometrial cancer. Patel *et al* studied tumor-adjacent and tumor endometrial tissue (Patel *et al* 2011), demonstrating the ability to distinguish between glandular epithelium, stroma and the lumen based on chemometrics analysis. A recent *in vitro* study by Borel *et al* using spectrally-resolved confocal Raman microscopy showed 92% sensitivity and 85% specificity to distinguish between normal and malignantly-transformed ovarian epithelial cells, a model for high-grade serous ovarian cancer (Borel *et al* 2015).

**Prostate**

The diagnosis and grading of prostate cancer play a key role in the patient survival. Different treatments such as active surveillance, radical prostatectomy (RP) and non-surgical therapy (radio-, chemo- or hormone-therapy) are available. For each patient, treatment is chosen depending on factors determined from histology of biopsies, the prostate specific antigen...
(PSA) level, clinical stage and patient condition. In the case of RP, a second diagnosis is established on tissue coming from the whole prostate and in 30% of cases differences between the two diagnoses are observed (Treurniet et al 2014). Moreover, the techniques are time consuming and in some cases diagnosis can be difficult to establish due to unusual cancer tissue morphology. There is, therefore, a need for new methods of prostate cancer detection that can supplement current techniques. To date, there have been no reports of RS for in vivo detection of prostate cancer. However, several studies have reported results using in vitro cell lines or ex vivo prostate samples. Wang et al used Raman micro-spectroscopy on two different cell lines (LNCaP and C4-2) and identified castration-resistant prostate cancer with a sensitivity and specificity of 88% and 88%, respectively (Wang et al 2013c). Crow et al similarly characterized different prostate adenocarcinoma cell lines (LNCaP, PCa2b, DU145 and PC-3) having varying aggressiveness, yielding 98% sensitivity and 99% specificity (Crow et al 2005a). The same group performed an ex vivo study with a fiberoptic Raman system, suitable for in vivo measurements, achieving an accuracy of 86% for cancer detection in 38 snap-frozen prostates taken at prostatectomy (Crow et al 2005b) and an accuracy of 89% on snap-frozen biopsies (Crow et al 2003).

Bladder

Transurethral resection of bladder tumors (TURBT) is the first line of cancer diagnosis, grading and treatment (removal of visible lesions). However, a recurrence rate of 50% is observed within 18 months (Allard et al 1998), because of the presence of small neoplastic lesions (carcinoma in situ, papillary lesion) that are overlooked during TURBT. RS has shown the potential to distinguish bladder cancer from normal bladder tissue: an in vitro study on 75 snap-frozen bladder samples differentiated between normal bladder, cystitis and transitional cell carcinoma/carcinoma in situ with sensitivity and specificity >90% using Raman micro-spectroscopy with 10s acquisition time per spectrum (Crow et al 2004); the same group achieved 84% accuracy on 29 snap-frozen bladders (Crow et al 2005b), while a similar study showed 94% sensitivity and 92% specificity in discriminating between non-tumor and tumor in snap-frozen tissue, with 20s per spectrum acquisition time (de Jong et al 2006). Finally, Draga et al used RS for in vivo detection of bladder cancer during TURBT and were able to distinguish cancer lesions from benign tissue with a 85% sensitivity and 79% specificity, and collection times of 1–5 s (Draga et al 2010). They also reported an increase in the intensity of specific amino acid peaks, likely associated with increased stage-dependent DNA concentration.

Oral

Oral cancers include lesions of the tongue, hard palate and floor of the mouth. Surgery is often used to remove the bulk of the tumor before chemotherapy or radiotherapy. This treatment is associated with high morbidity, as large surgical margins (<1 cm) are taken. Several optical techniques have been investigated to diagnose oral cancer and guide surgical resection (Singh et al 2015).

Early animal studies used RS on chemically-induced dysplasia on rat palate, giving 78% sensitivity and 93% specificity for low-grade dysplasia and 100% sensitivity and specificity for high-grade dysplasia, but requiring acquisition times of 100 s for the training data and 10 s for the test data (Bakker Schut et al 2000). Guze et al acquired in vivo Raman spectra from 7 different sites in the mouth of 51 healthy patients with 1s integration time in the 1500–3100 cm⁻¹ band (Guze et al 2009). Classification yielded mixed results, with the
accuracy ranging from 60% to 100%. The best results were obtained in the 2800–3100 cm\(^{-1}\) region. This study provided a baseline for future \textit{in vivo} applications. Singh \textit{et al} published a study comparing RS in tumor and healthy tissue from 50 patients, yielding 86% sensitivity and 83% specificity (Singh \textit{et al} 2012b). This was followed by a study in 104 patients, acquiring 861 spectra from normal, premalignant and malignant sites (Singh \textit{et al} 2012a). They compared results with healthy subjects and pre-malignant sites, with a classification accuracy of 95%. Krishna \textit{et al} used \textit{in vivo} RS for the detection of normal oral mucosa, squamous cell carcinoma, submucosa fibrosis and leukoplakia achieving accuracies of 85%, 89%, 85%, and 82%, respectively (Krishna \textit{et al} 2014). Distinguishing normal from abnormal (pooling the other 3 types) resulted in 94% sensitivity and specificity.

\textbf{Gastrointestinal (GI) tract}

The GI tract consists of a number of organs including the pancreas, esophagus, stomach and intestines. For many GI cancers, early detection is critical to reducing mortality rates. Due to the prevalence of endoscopy for GI diagnosis, a number of groups have developed endoscopic RS systems. Long acquisition times and the need to miniaturize fiber-optic components are some of the primary challenges for implementing RS in this manner. While there has been a great deal of \textit{in vitro} work done (Teh \textit{et al} 2008, Widjaja \textit{et al} 2008, Kawabata \textit{et al} 2011), we will focus on \textit{in vivo} studies: Bergholt \textit{et al} provides a comprehensive review of both \textit{ex vivo} and \textit{in vivo} work in the GI tract (Bergholt 2013). Shim \textit{et al} used RS \textit{in vivo} to differentiate adenomatous and hyperplastic polyps in the colon with 100% sensitivity and 89% specificity, using 5 s acquisition time (Shim \textit{et al} 2000). Bergholt \textit{et al} have developed \textit{in vivo} RS endoscopy systems with <1 s acquisition times, and achieved 85% sensitivity and 96% specificity to detect adenocarcinoma in gastric tissue \textit{in vivo} during endoscopy (Bergholt \textit{et al} 2013). This group has also distinguished between normal mucosa, benign and malignant ulcerous lesions in the stomach with sensitivities of 91%, 85%, 82% and specificities of 94%, 95%, 95%, respectively (Bergholt \textit{et al} 2010). The Wilson group assessed both \textit{in vivo} and \textit{ex vivo} tissue in the colon for distinguishing adenomatous and hyperplastic polyps, with 100% sensitivity and 89% specificity for \textit{in vivo} samples, with 30 s acquisition time (Molckovsky \textit{et al} 2003).

They also reported detecting high-grade dysplasia in Barrett’s esophagus with ~90% sensitivity and specificity (Wilson 2007) (figure 4). Combining high-wavenumber with fingerprint RS \textit{in vivo}, Wang \textit{et al} detected esophageal squamous cell carcinoma with 97% sensitivity and specificity, using acquisition times of 0.1–0.5 s (Wang \textit{et al} 2015). This body of \textit{in vivo} work that has achieved high classification performance indicates the diagnostic potential for endoscopic RS in the GI tract. This may also be applicable in lung cancer; a recent study used endoscopic RS \textit{in vivo} for early detection of lung cancer. In 280 samples from 80 patients, they achieved 90% sensitivity and 65% specificity for the detection of high grade dysplasia and malignant lung lesions, with an acquisition time of 1 s (McGregor \textit{et al} 2016).

\textbf{Biofluids}

RS has also seen increased use to measure biofluids such as blood, urine and saliva for the detection of disease and as a prognostic indicator for treatment monitoring. Biofluids are rich in chemical information, motivating the application of RS. Many standard-of-care tests for disease lack the desired specificity for effective screening. For example, the prostate-specific antigen (PSA) test is commonly used for prostate cancer screening but only ~25% of men who undergo biopsy due to elevated PSA level have prostate cancer (Barry 2001). RS has potential for a
highly-specific, low-cost and non-invasive optical test to complement or replace existing procedures. Austin et al have written a comprehensive review of biofluid RS (Austin et al 2016); here, we will highlight some of the more recent developments for screening and detection.

It is common to use SERS for the analysis of biofluids, due to the full surface coverage and improved signal detection. Feng et al used SERS with an integration time of 1s to analyze blood plasma for colorectal cancer and adenomatous polyps, achieving 86% sensitivity and 80% specificity (Feng et al 2015b). This group also recently used SERS on saliva to differentiate healthy subjects from subjects with benign and malignant breast tumors, reporting sensitivities of 75%, 72% and 74% and specificities 94%, 81% and 86% (Feng et al 2015a). SERS with silver nanoparticles has been used for the detection of PSA in serum, giving diagnostic accuracy of 98% (93 prostate cancer patients, 68 healthy patients) and acquisition time of 10s (Li et al 2014). Lin et al assessed blood serum using SERS in 38 patients with colorectal cancer and 45 healthy subjects, achieving 97% sensitivity and 100% specificity with 10s integration time (Lin et al 2011). Elumalai et al achieved 99% sensitivity and 87% specificity for detection of oral cancer using urine-based RS, with 90s integration time (Elumalai et al 2015). RS has been used on blood serum from 11 patients with breast cancer and 12 healthy subjects, achieving 97% sensitivity and 78% specificity with a 10s acquisition time (Pichardo-Molina et al 2007). It has also been used for Alzheimer’s disease diagnosis based on blood serum, with >95% sensitivity and specificity, using two 10s acquisitions per spectrum (Ryzhikova et al 2015). While the reported sensitivities and specificities of biofluid-based RS detection are still somewhat limited for many applications, there is a great deal of potential for improvements to current standard-of-care cancer screening, particularly to supplement existing diagnostic methods.
Multimodal imaging with Raman spectroscopy

The molecular information from RS can be combined with complementary information from other modalities to improve diagnosis or for treatment planning and guidance. While in some cases the achieved in vivo sensitivities and specificities with RS may be sufficient for clinical translation, there are many diagnostic applications that could benefit from complementary biomarkers. For example, metabolic (tissue fluorescence) and morphological (optical coherence tomography) information can serve as quantitative diagnostic biomarkers, and potentially improve the sensitivity and specificity. For many applications, such as brain tumor resection, recurrence is common due to residual cancer remaining after surgery, and so improvements to diagnostic performance are particularly important in approaching complete resection.

Bergholt et al combined RS with near-infrared autofluorescence in vivo for endoscopic detection of gastric cancer, resulting in 98% sensitivity and 92% specificity that were better than RS alone (Bergholt et al 2011). Zakharov et al also used autofluorescence imaging combined with RS for in vivo detection of malignant skin melanoma, achieving 89% sensitivity and 87% specificity (Zakharov et al 2014b). Combined RS and optical coherence tomography (OCT) has been demonstrated ex vivo in rodents and in vivo in human skin, with 30 s acquisition time for RS (figure 5) (Patil et al 2011). This is a potent combination as OCT provides tissue microstructural information but lacks the molecular information that RS provides. RS-OCT was used to distinguish normal colon from colonic adenocarcinoma, achieving 94% sensitivity and specificity with a 5 s acquisition time for RS, significantly outperforming either modality alone (Ashok et al 2013). RS-OCT has also been used on the skin for the detection of malignant melanoma with 89% sensitivity and 88% specificity, adenocarcinoma with 100% sensitivity and 82% specificity, and squamous cell carcinoma with 91% sensitivity and 78% specificity (Zakharov et al 2014a). CARS has been used with two-photon excited fluorescence to image
myelinated fibers in the mouse brain, helping to identify fiber bundles (Fu et al. 2008). Wang et al. demonstrated that reflectance confocal microscopy and multiphoton microscopy imaging can be used with RS to measure micro-structures of human skin in vivo, with the ability to determine blood flow velocity and blood glucose level (Wang et al. 2013a). A great deal of progress has been made for the use of multimodal imaging and spectroscopy to complement the molecular detection of RS.

**Challenges for clinical implementation**

There are a number of factors that make RS well suited to clinical translation as compared to other imaging techniques, including the lack of exogenous contrast agents, thereby enabling easier integration into the clinical workflow and reduced regulatory barriers. Recent advances in instrumentation and spectral analysis have allowed for sub-second acquisition times and high diagnostic performance (Saar et al. 2010, Keller et al. 2011, Ji et al. 2013, Liao et al. 2015, Jermyn et al. 2015). The ability to make near real-time measurements is important for in vivo interventional applications. In addition, Raman-based imaging and spectroscopy systems can be made portable and clinically ergonomic. Thus, a number of hand-held tools have been developed specifically for in vivo Raman spectroscopy (Freudiger et al. 2014, Jermyn et al. 2015).

Despite these compelling advantages, several challenges have limited clinical adoption of RS. The Raman signal is typically dominated by elastic light scattering and intrinsic tissue autofluorescence. Techniques such as SERS and Coherent RS aim to address this issue with instrumentation changes that can substantially improve the measured signal-to-background ratio. However, these advantages often come at the cost of increased acquisition time, complex instrumentation cost and other factors that impede clinical translation. Not only is the low signal-to-noise a challenge, but the diagnostically relevant signal is also spread across the spectral bands. This limits the use of standard univariate analysis methods. Sophisticated machine-learning techniques have been developed that can take advantage of the rich spectral information.

Choosing the most practical and accurate technology often involves trade-offs between field-of-view, spatial resolution, spectral resolution and acquisition time. There is frequently a trade-off between spatial and spectral resolution, and more detailed studies evaluating the most relevant information for different applications are required.

**Signal detection**

One of the primary issues in using RS for clinical diagnostics is the trade-off between acquired molecular signals and the time needed for acquisition. The development of CARS, SRS and SERS were, at least in part, aimed at improving the signal-to-noise ratio and recent developments have significantly reduced the acquisition time. CARS and SRS have achieved video-rate imaging speed and can use frequency multiplexing. They are also able to achieve very high spatial resolution ~1 µm (Saar et al. 2010, Meyer et al. 2011). However, many of these systems are complex and difficult for clinical implementation due to issues with efficient light delivery and collection, and miniaturization of optical components (Latka et al. 2013). Moreover, while acquisition times are impressively fast for point-detection, spectroscopic imaging still takes substantial time. For many diagnostic applications, imaging is more appropriate, especially when imaging dynamic systems in vivo. Raman techniques such as SORS can increase the depth at which Raman signal can be measured (Stone et al. 2007). This is applicable in clinical
situations where subsurface disease detection is important, such as margin assessment for cancer: the ability to detect cancer cells beneath tissue at the periphery of a tumor could be crucial for many types of cancer.

Some of the technical challenges for Raman system development come from the nature of the diagnostic applications of interest. For endoscopic systems and some probe-based systems, size limitations require component miniaturization that can affect the signal detection: see the section on Raman techniques for disease detection and table 1. There are limitations on the permissible laser exposure in order to avoid tissue damage, based on ANSI Z136.1 laser safety standards. A confounding factor for Raman signal detection is the difficulty of making quantitative measurements of tissue composition in vivo, due to the number and heterogeneity of molecular species and the confounding effects of optical absorption and (elastic) scattering. Chemometrics-based analysis can be used to spectrally decompose the molecular components from RS, but must be used in conjunction with reliable calibration and data-processing procedures in order to preserve quantitative information, and often have difficulty with the large number of molecular species for the acquired Raman signal. Validation methods for such algorithms are also needed to ensure that quantitative information retrieved from the tissue spectra is reliable. More robust quantification in RS would benefit screening and intraoperative methods.

Tissue volume sampling and preservation

One of the barriers in moving Raman techniques forward for clinical applications is tissue sampling for validation studies, correlating in vivo Raman measurements with ex vivo histopathology as the ‘gold standard’. However, it is often difficult in practice to ensure that the tissue sample (e.g. biopsy or resection) is co-registered exactly with that of the Raman measurement. The confounding factors are: (i) the Raman instrument characteristics, particularly the probe tip size and geometry, (ii) the tissue absorption and elastic scattering that determine the effective sampling depth and volume, (iii) the biopsy tool (lateral extent and depth of tissue cutting) and how this is co-located with the Raman probe, and (iv) handling, preservation and orientation of the biopsy for histopathologic sectioning. Most in vivo studies have focused on taking superficial tissue biopsies, since the majority of the signal will come from a few hundred microns from the surface except for SORS applications where sampling depth can be up to several millimeters.

An added issue is that, even if the tissue biopsy and the effective Raman sampling volume are spatially co-located, there is often marked heterogeneity within the tissue sample, e.g. it contains both tumor and normal tissues or several different grades of tumor. Generally, for clinical pathology, the highest grade is the most important for determining treatment and this may be all that is reported, even if it comprises only a tiny fraction of the tissue. In this case, the RS result would be considered as false negative relative to the histopathology since this highest-grade tissue would contribute only a small fraction of the Raman signal. In order to address this, at least in part, the approximate extent of disease in the biopsy sample can be determined by the pathologist in order to improve the correlation accuracy.

The best alternative to making in vivo Raman measurements is the use of fresh tissue samples, kept on ice or in cold saline after removal, and performing the measurements within a few hours. For longer intervals, tissues are either snap frozen or formalin fixed and paraffin embedded (FFPE). FFPE is commonly used by biobanks for long-term preservation. Care is required in RS on FFPE tissues since there may be contributions to the signal from some of the chemicals used in the FFPE and dewaxing (paraffin removal) processes. Faoláin et al
found deterioration in the spectra of frozen tissue compared to fresh tissue, with some spectral peaks vanishing (Faoláin et al 2005a, 2005b). Deterioration has also been observed in FFPE tissues before and after dewaxing, as well as the appearance of new peaks (Faoláin et al 2005b). Dewaxing and FFPE can introduce optical artifacts and diminish the spectral differences between normal and malignant tissue (Crow et al 2002, Krishna et al 2005). Formalin is not a major pollutant of Raman spectra, but affects the cellular lipid and protein content, as shown on in vitro cell culture lines (Mariani et al 2009).

**Ambient light effects**

External sources of light can impair the ability to properly measure Raman spectra in a clinical setting. Endoscopic applications are less prone to ambient light artifacts as open-field surgical applications where there are operating room lights, LCD screens, daylight leakage and for some applications infrared navigational tracking units (Desroches et al 2015). It is, therefore, important to determine the permissible ambient light level during RS acquisition and, if necessary eliminate this at least temporarily. For steady ambient light, it may be possible either to modulate the RS light source and detect synchronously with it (Dochow et al 2012), to correct for the contribution in the spectral analysis (Jermyn et al 2016) (see Tissue classification below) or normalize the signal using a background measurement acquired with the tissue excitation sources turned off.

**Spectral pre-processing**

Due to the many common vibrational and rotational modes of the molecular constituents of cells and tissues, the entire Raman spectrum must typically be analyzed to fully utilize all the diagnostic information, rather than looking simply at changes in individual peaks. Pre-processing steps are usually required first to reduce noise, remove background, and normalize the spectra to enable comparisons between data sets. The inelastic Raman scattering is dominated by elastic scattering and intrinsic fluorescence, and so stochastic noise is a limiting factor for Raman spectra (Ramírez-Elías et al 2012, Van de Sompel et al 2013, Smulko et al 2014). Fourier filtering can be applied to suppress this noise, which is in the high-frequency range. Alternatively, smoothing functions such as moving average can be used. Both approaches can sometimes remove real Raman signals due to overlapping spectral frequencies and so must be applied with care. Cosmic rays can also cause spurious noise, but are rare and produce distinct features, so that they can easily be detected and removed with repeated measurements (Zhang and Henson 2007, Li and Dai 2011, Schulze and Turner 2013).

Background removal can involve the removal of both the measured ambient signal without laser excitation and the non-Raman signal from tissue autofluorescence. The former can simply be determined immediately before the true measurement is taken, and then subtracted. The latter is more problematic, since, even in the near-infrared, the intrinsic fluorescence of tissue dominates the measured spectrum. However, the fluorescence spectrum is usually broad and relatively smooth, so that baseline-correction methods can be used to remove it (Schulze et al 2005), as long as the instrument has a large enough dynamic range. Fourier filtering and polynomial curve fitting are also commonly used for baseline correction (Schulze et al 2005, Zhao et al 2007). Instrumentation changes can also be implemented to assist in background removal, including time gating to remove the relatively long-lived fluorescence (Morris et al 2005) or subtracting the spectra taken at two slightly different excitation wavelengths where the fluorescence spectra are nearly identical but the Raman peaks are shifted (Shreve et al 1992).
After background removal, normalization can be applied, either by normalizing the whole spectrum or normalizing restricted spectral bands in the case where particular peaks or bands are analyzed individually. In both cases the spectral signal is divided by a metric such as variance, standard deviation, area under the curve, or specific band heights. A standard reference material can also be measured to calibrate the spectra, ensuring that observed spectral features can be specifically associated with those of known Raman-active molecules. Normalization also risks diminishing some of the true spectral changes with disease, so that their impact should be assessed for the particular application.

**Spectral analysis and feature extraction**

For *in vivo* applications, analyzing a single Raman band often does not capture enough information to be of diagnostic use. There are a large number of Raman-active molecules in tissue, each with its own Raman fingerprint, so that the measured spectrum is a weighted combination of these contributions and the pertinent information is spread across the spectrum. Moreover, disease-related changes may not be expressed predictably in a single molecular species. Hence, analysis methods utilizing all of the available spectral information are usually required. Some of these methods are applied directly to the full spectra, but it is also possible to first use feature extraction algorithms to identify the most important spectral bands or other spectral characteristics, thereby selecting a subset of the data or creating new features. In either case, the goal is to eliminate redundant or irrelevant information for subsequent classification. This has multiple benefits, including reducing the computational complexity, improving interpretability and avoiding overfitting of the data.

A frequently used feature-extraction method is principal component analysis (PCA) (Li et al 2013), using orthogonal linear combinations to create uncorrelated variables, known as principle components (PCs). The PCs are created to maximize the variance in the data, with the added constraint of orthogonality to ensure that the PCs are independent of one another. The lowest-order PCs carry the greatest information and it is important to restrict the number of PCs used in the subsequent classification algorithms. Note that PCA is unsupervised, in the sense that the true classes are not given as inputs (e.g. cancer versus normal tissue), so that some of the information may not be related to the diagnostic objective but are due to variance caused by other factors. There are a number of other feature selection and extraction methods that can be used in RS, such as discrete cosine transform, multifactor dimensionality reduction, non-negative matrix factorization, correlation feature selection, and Fisher-based feature selection (Ahmed et al 1974, Hall and Smith 1999, Ritchie et al 2003, Dhillon and Sra 2005, Fenn et al 2013, Zhang et al 2015), each with its own advantages and limitations.

**Tissue classification**

Supervised machine-learning uses *training data* to create a model for making predictions. In the context of RS, these data consist of a set of independent spectra for which the biological or clinical status is known. The desired predictions are often called *classes* and typically concern the presence of disease, diagnostic grade or other clinical information such as treatment outcome. A common example would be classifying cancer from normal tissue compared to histopathological ‘truth’. Either the full Raman spectra or specific features may be used as training data. The classification algorithm is then trained on these data and subsequently used to make predictions about *test data* where the classification is not known a priori or is treated as unknown for validation purposes. Linear discriminant analysis (LDA) is a common
classification algorithm (Li et al 2013) which is similar to PCA in that it uses linear combinations to evaluate the data. However, unlike PCA, which does not use class information and so is unsupervised, LDA seeks to maximize the data variation between different known classes. A second common classification approach is to use support vector machines (SVM) that operate by finding the hyperplane which best separates the data into the classes, based on maximizing the separation between data from different classes (Widjaja et al 2008). Artificial neural networks (ANN) is a classification method inspired by the functioning of neurons in the brain. The ‘neurons’ are arranged in layers, with the first layer connected to the data (here, the Raman spectra), the last layer connected to the class output (e.g. cancer or normal tissue) and each of several intermediate layers connected to its neighboring layers below and above (figure 6). The connections, represented by lines of different colors in the figure, carry weights and each neuron applies a non-linear transformation to the weighted sum of its inputs. The optimum weights are determined from the training data. ANN is very useful for modeling complex non-linear interactions, making it appropriate for use with Raman spectra (Gniadecka et al 2004, Sigurdsson et al 2004, Jermyn et al 2016).

The boosted-trees method has also been used for RS classification (Jermyn et al 2015) by constructing an ensemble of data-based decision trees, with each tree operating on the residual of the previous tree in order to distinguish between classes. The random-forests technique (Seng Khoon Teh 2009) similarly constructs an ensemble of decision trees and then employs their mode or mean for classification. This avoids the problem of overfitting common to the use of decision trees.

Thus, there are a variety of classification algorithms available and the appropriate choice will depend on the problem and data-specific details. However, poor data collection, preprocessing or feature extraction may result in sub-optimal classification regardless of the choice of algorithm. When using other imaging or spectroscopy modalities together with Raman measurements, it is possible to combine information for the purposes of classification. Hierarchical or sequential classification can be used either to apply classification methods to predicted values from other classifiers (El-Shishiny et al 1989, Kurzynski 1989, Pudil et al

Figure 6. Conceptual diagram of a classification model using artificial neural networks to determine tissue diagnostic information from Raman spectra.
1992) or to utilize voting-based techniques that apply weights to the predicted outcomes from different classifiers. In fact, ensembles of classifiers can be used to improve generalization and classification performance even when operating with a single modality. Some advanced techniques construct large numbers of ensemble classifiers and then apply methods such as random forests to maximize the performance gained while avoiding the vulnerability to noise and overfitting found with individual classifiers.

Unsupervised techniques such as vertex component analysis (VCA), hierarchical cluster analysis (HCA), and singular value decomposition (SVD) can separate Raman spectra without using class information, and are frequently used with Raman imaging approaches. VCA and SVD attempt to model spectra as linear combinations of component spectra (Khmaladze et al 2014, Tabarangao et al 2015), which in the ideal case would correspond to relative contributions from different molecular species, while HCA iteratively clusters spectra based on computed distances (Bonifacio et al 2015). Finally, regression is a type of supervised machine-learning where, instead of predicting discrete classes, the desired output is one or more variables and the input and output variables are related using machine-learning. This can be useful for determining concentrations of different molecular species in tissue based on Raman spectra.

Validation of classification performance

Classification techniques may perform well on training data but then perform poorly on the independent test data. A common cause is overfitting the model to the data, i.e. effectively fitting to the noise in the training data. Conversely, the model needs to be complex enough to distinguish robustly between the classes. Cross-validation can be used to assess the performance of a particular model, determining metrics such as accuracy, sensitivity, specificity and receiver operating characteristic curve (ROC) analysis. In this, the data are (randomly) partitioned into separate training and validation sets. The former is used to train the algorithm which is then tested on the validation set. This process of partitioning, training and testing is repeated multiple times, using methods such as leave-one-out cross validation (LOOCV) or k-fold cross validation (Gautam et al 2015) to select the data sub-sets. A caveat is to ensure that inter-dependent spectra never appear in both the training and testing datasets: e.g. if a single tissue sample is measured twice, the spectra will be highly correlated. It is important to ensure sufficient sample sizes when using cross-validation techniques for statistical robustness. When making the choice of classification algorithm based on the performance under cross validation, bias may occur. Thus, if multiple algorithms are tested, some may perform well simply by chance (based specifically on the data used). To provide further robustness against overfitting and selection bias, a hold-out dataset can be used that is separate from both the training and validation data. This set is then used after the optimal classification algorithm is selected and provides a means of assessing the performance in terms of generalization to new data. Moreover, it is possible in certain classification algorithms to simplify the model using pruning or other methods in which the least significant nodes, trees or other elements are removed. For iterative classification algorithms, providing early-stopping criteria can also help prevent overfitting.

Discussion

Vibrational spectroscopy can yield a wealth of molecular information from biological tissue, which can be used to discover and quantify new intrinsic biomarkers associated with disease. Raman spectroscopy in particular is seeing increased clinical investigation as the
instrumentation and spectral analysis techniques improve. This family of techniques has a number of implementation challenges, requiring interdisciplinary contributions from clinicians and biologists as well as engineers, mathematicians and physicists. As our understanding of the molecular basis of disease improves, this information can be incorporated into the optimal design of Raman systems. Disease progression is typically associated with molecular changes, motivating the use of RS for detection, as well as planning and guiding treatment. RS can be used in place of or in combination with existing clinical practices such as biopsy, histopathology and radiologically-based surgical guidance.

Signal detection remains one of the primary challenges facing the use of RS for clinical diagnostics. The contribution from intrinsic auto-fluorescence background in tissue often dominates the relatively small Raman signal, and other factors such as noise, instrument response, and ambient light further confound the problem. While sophisticated spectral analysis and pre-processing techniques can mitigate some of the confounding factors, signal detection is still the key limitation. Many approaches such as CARS, SRS, and SERS aim to address this issue, and evaluating the trade-offs of these approaches in the context of different oncology applications will be critical to enabling clinical translation, particularly regarding acquisition time and spectral/spatial resolution.

Promising in vivo results across a range of applications indicate the ability to overcome the technical challenges facing the use of RS in the real clinical environment and motivate the need for further advances. In particular, substantial advances have been made in cancer diagnostics, with many in vivo trials achieving >90% sensitivity and specificity, often with acquisition times <1 s. The recent surge in machine-learning research has benefitted RS, making available sophisticated classification algorithms that can optimally utilize the spectral information. While surgical guidance is one of the primary avenues for clinical use of RS, the increased portability, convenience and affordability of RS technology also makes it increasingly viable for screening and point-of-care, extending Raman techniques into a wider clinical user community than has been the case to date.

Acknowledgments

This work is supported by the Fonds de recherche du Québec—Nature et technologies (FRQNT), the Discovery Grant program from Natural Sciences and Engineering Research Council of Canada (NSERC), the Collaborative Health Research Program (CIHR and NSERC), and fellowships from the Groupe de Recherche en Sciences et Technologies biomédicales (GRSTB), Banque Nationale, and FRQNT.

Competing interests

KP and FL are co-founders of ODS Medical Inc., a medical device company that seeks to commercialize Raman spectroscopy systems for real-time detection of tissue abnormalities.

References

Austin L A, Osseiran S and Evans C L 2016 Raman technologies in cancer diagnostics Analyst 141 476–503
Bergholt M S 2013 Raman endoscopy for objective diagnosis of early cancer in the gastrointestinal system J. Gastrointest. Dig. Syst. 81 008
Boca-Farcau S, Potara M, Simon T, Juhem A, Baldeck P and Astilean S 2014 Folic acid-conjugated, SERS-labeled silver nanotriangles for multimodal detection and targeted photothermal treatment on human ovarian cancer cells Mol. Pharm. 11 391–9
Cloyd K L et al 2012 Characterization of porcine aortic valvular interstitial cell ‘calcified’ nodules PloS One 7 e48154
Fenn M B, Pappu V, Georgeiv P G and Pardalos P M 2013 Raman spectroscopy utilizing Fisher-based feature selection combined with Support Vector Machines for the characterization of breast cell lines J. Raman Spectrosc. 44 939–48


Goodenberger M L and Jenkins R B 2012 Genetics of adult glioma Cancer Genet. 205 613–21


Je et al 2013 Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy Sci. Transl. Med. 5 201ra119


Li S and Dai L 2011 An improved algorithm to remove cosmic spikes in Raman spectra for online monitoring *Appl. Spectrosc.* **65** 1300–6
Li Y *et al* 2013 Micro-Raman spectroscopy study of cancerous and normal nasopharyngeal tissues *J. Biomed. Opt.* **18** 27003
Lord R C and Yu N 1970 Laser-excited Raman spectroscopy of biomolecules *J. Mol. Biol.* **50** 509–24
Lui H, Zhao J, McLean D and Zeng H 2012 Real-time raman spectroscopy for *in vivo* skin cancer diagnosis *Cancer Res.* **72** 2491–500


