Aluminium foil as an alternative substrate for the spectroscopic interrogation of endometrial cancer

Maria Paraskevaidi[a,1], Camilo L.M. Morais[a,b], Olivia Raglan[c], Kássio M.G. Lima[b], Evangelos Paraskevaidis[d], Pierre L. Martin-Hirsch[e], Maria Kyrgiou[c] and Francis L. Martin[a,1]

[a]*School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK*

[b]*Institute of Chemistry, Biological Chemistry and Chemometrics, Federal University of Rio Grande do Norte, Natal 59072-970, Brazil*

[c]*Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, London W12 0HS, UK*

[d]*Department of Obstetrics and Gynaecology, University Hospital of Ioannina, Ioannina 45500, Greece*

[e]*Department of Obstetrics and Gynaecology, Lancashire Teaching Hospitals NHS Foundation, Sharoe Green Unit, Fullwood, Preston PR2 9HT, UK*

[1]To whom correspondence should be addressed. Email: mparaskevaidi@uclan.ac.uk or flmartin@uclan.ac.uk

Abstract

Biospectroscopy has the potential to investigate and characterise biological samples and could, therefore, be utilised to diagnose various diseases in a clinical environment. An important consideration in spectrochemical studies is the cost-effectiveness of the substrate used to support the sample, as high expense would limit their translation into clinic. In this paper, the performance of low-cost aluminium (Al) foil substrates was compared with the commonly used low-emissivity (low-E) slides. Attenuated total reflection-Fourier transform infrared (ATR-FTIR) spectroscopy was used to analyse blood plasma and serum samples from women with endometrial cancer and healthy controls. The two populations were differentiated using principal component analysis with support vector machines (PCA-SVM) with 100% sensitivity in plasma samples (endometrial cancer=70; healthy controls=15) using both Al foil and low-E slides as substrates. The same sensitivity results (100%) were achieved for serum samples (endometrial cancer=60; healthy controls=15). Specificity was found higher using Al foil (90%) in comparison to low-E slides (85%) and lower using Al foil (70%) in comparison to low-E slides in serum samples. The establishment of Al foil as low-cost and highly-performing substrate would pave the way for large-scale, multi-centre studies and potentially for routine clinical use.

Introduction

Vibrational spectroscopy is increasingly utilised in biomedical research as a valuable tool in disease investigation. Allowing the analysis of a variety of biological samples, such as cells, tissues and biofluids, this spectrochemical analysis has a bright future ahead, not only in scientific/laboratory research but also in clinical practice. The key factor that renders this analytical method a perfect diagnostic tool, in comparison to other molecular methods, is its non-destructive, cost-effective and label-free nature. Over the years, infrared (IR) and Raman spectroscopic techniques have been employed to study a number of different diseases like cancer, neurological diseases, prenatal disorders and many others [1-10]. Within the field of disease investigation, spectroscopy has the potential to diagnose and monitor a disease, while at the same time assessment of surgical margins of a tumour or determination of the subtype of a disease is also feasible.

Most spectroscopic studies so far, with only a few exceptions [8, 10, 11], have included a limited number of subjects which appears to be an important limitation for the establishment of the method and its migration into clinics [12-14]. Standardisation and validation of methods should be performed in large clinical trials for more robust and trustworthy results. A further issue that limits the ability for clinical implementation relates to experimental methodology. Specifically, inconsistencies in the pre-analytical stages of sample collection and preparation to spectral collection and data analysis. A fundamental factor of the analytical procedure is the use of the correct substrate in order to avoid non-biological interference from the substrate in use. Unfortunately, the majority of the available, "featureless" substrates are high-cost [15, 16], something which prevents their use in large scale studies and routine analysis. Previous studies have even developed data correction algorithms to remove the substrate's signal after the collection of the raw spectra [16-18].

Different types of substrates are selected depending on the spectroscopic technique used each time (*e.g.*, IR or Raman spectroscopy), as well as on the chosen sampling mode [*e.g.*, transmission IR, transflection IR or attenuated total reflection (ATR)]. Namely, some of the substrates that have been used for IR and Raman spectroscopy over the years include barium fluoride (BaF_2), calcium fluoride (CaF_2), zinc selenide (ZnSe), gold-coated (Au), silver or silver-coated (Ag), fused silica (SiO_2) and low-emissivity (low-E) slides [19, 20]. However, due to their expense, efforts are being made to introduce novel, low-cost substrates that would facilitate the analysis of hundreds, even thousands, of samples cost-effectively. Glass substrates are routinely used in medical laboratories and hospitals for preparation of analysis of various types of biological samples; however, glass has been found unsuitable for spectroscopy as it generates background signal and distorts the biological information coming from the samples [19]. Therefore, an ideal approach would be to take advantage of the extremely cost-effective glass slides by covering them with a metallic surface that would eliminate any background noise. Previous proof-of-concept studies have been conducted showing aluminium (Al) foil's potential as a suitable substrate [11, 16, 21]. A robust and inexpensive substrate for both IR and Raman spectroscopic methods would be extremely beneficial. However, there has been no conclusive study comparing its effect on diagnostic accuracy with other, widely used substrates.

In this study, we used ATR-FTIR spectroscopy to explore whether Al foil could be an appropriate substrate for spectroscopic investigations. ATR-FTIR uses an internal reflection element (IRE) with a high refractive index to direct the beam to the sample; an evanescent wave is created, penetrating the sample by a few microns in order to derive its chemical information [22]. A commonly used substrate for ATR-FTIR measurements is the low-E slide, which has been effectively used in numerous biological studies in the past [23-25]. Therefore, we compared our results from the low-E slides with those from Al foil slides to assess the

performance of the latter with regard to the diagnostic accuracy. For the purpose of this piece of work, we analysed blood samples from women with endometrial cancer, as well as from benign cases used as controls. Endometrial cancer develops in the endometrium (*i.e.*, inner lining of the uterus) and is the fourth most common gynaecological cancer in the developing world, with an increasing incidence in postmenopausal women; in 2012 alone, 319,000 new cases were diagnosed worldwide [26]. Although symptoms of endometrial cancer develop relatively early, which allows "timely" diagnosis and early intervention, a more objective, less expensive and non-invasive method of diagnosing this type of cancer is highly desirable and clinically indicated. Currently, a diagnosis is based on microscopic histological examination of endometrial tissue, which is dependent on subjective interpretation, therefore allowing human error.

Materials and Methods

Blood plasma and serum analysis

The collection of all samples for this study was approved by the institutional review board at Imperial College Healthcare NHS Trust (tissue bank sub-collection number GYN/HG/13-020). All patients provided informed consent for use of their samples in this study. This study included age-matched cohorts; plasma samples were available for 70 endometrial cancer patients and 15 non-cancer individuals used as controls; serum samples were available for 60 endometrial cancer patients and 15 controls. At time of diagnosis, patients were not receiving any medications such as Tamoxifen treatments which might affect the outcomes. Also women who had hyperplasia or hypertension have been excluded. Both blood plasma and serum samples were collected and stored at -80°C until analysis; prior to spectroscopic interrogation, the samples were left to defrost at room temperature before 50 μL of each were deposited on a substrate and left to air-dry for approximately 30 min. All of the samples were analysed in

duplicates using two different substrates: the IR-reflective glass slides (MirrIR Low-E slides, Kevley Technologies, USA) and cheap, microscope glass slides covered with Al foil. The latter were carefully flattened with the shiny side of the foil being exposed to achieve a greater level of reflectivity. Covering the slide with Al foil required ~30-45 seconds with one slide taking up to 3 different samples, rendering the slide preparation time insignificant.

Spectrochemical Analysis

All blood samples were analysed using a Tensor 27 FTIR spectrometer with Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). The sampling area, defined by the internal reflection element (IRE), which was a diamond crystal, was approximately 250 μm × 250 μm. The slide with the sample is placed onto a moving platform with the sample facing up; the platform is then moved upward to achieve good contact with the diamond crystal. Spectral resolution was 8 cm^{-1} with two times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000-400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal to noise ratio. A CCTV camera attachment was used to locate the area of interest and spectra were acquired from ten different locations to minimize bias. Also, in order to take into consideration the natural phenomenon of "coffee ring" effect, spectra were mainly collected from the periphery of each drop where the absorbance intensity was higher, as important components, such as proteins and nucleic acids, migrate towards the edge of the drop after drying [27]. The ATR crystal was cleaned with distilled water before moving to a different sample and a background spectrum was acquired to take into account any atmospheric changes.

Spectral data handling and analysis

All spectral information was converted to suitable files (.txt) before input to MATLAB (Mathworks, Natick, USA). Pre-processing and computational analysis of the data was performed using PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA) and

141 an in-house developed IRootLab toolbox (http://trevisanj.github.io/irootlab/). Pre-processing
142 of the acquired spectra is an essential step of all spectroscopic experiments and is used to
143 correct problems associated with spectral acquisition, instrumentation or even sample handling
144 before further multivariate analysis [28]. In this study, spectra were cut at the biochemical
145 fingerprint region (1800-900 cm^{-1}), rubberband baseline corrected and vector normalised.

146 The samples were divided into training (~70%), validation (~15%) and test (~15%) sets
147 on a patient basis before chemometric analysis, using the Kennard-Stone sample selection
148 algorithm [29]; all spectra collected for each individual were used for model construction. In total,
149 60 samples were used for training (n = 600 spectra), 12 for validation (n = 120 spectra) and 13
150 for test (n = 130 spectra) with plasma samples; and 53 for training (n = 530 spectra), 11 for
151 validation (n = 110 spectra) and 11 for test (n = 110 spectra) with serum samples. The training
152 set was used for model construction, the validation set for optimization of the number of
153 principal components and latent variables used, and the test set for final model evaluation.
154 Cross-validation venetian blinds (10 splits with 1 sample per split) was used for optimization
155 of support vector machines (SVM) parameters (cost, epsilon, gamma and number of support
156 vectors) in principal component analysis with support vector machines (PCA-SVM).

157 For the classification of endometrial cancer and non-cancer cases a number of
158 chemometric techniques was used, such as partial least squares discriminant analysis (PLS-
159 DA); and principal component analysis followed by linear discriminant analysis (PCA-LDA),
160 quadratic discriminant analysis (PCA-QDA) and support vector machines (PCA-SVM).

161 PLS-DA is one of the most known chemometric technique of supervised classification.
162 It is based on a linear classification model for which the classification criterion is obtained by
163 partial least squares (PLS) analysis [30]. For this, PLS is applied to the data reducing the original
164 variables (*e.g.*, wavenumbers) to a few number of latent variables (LVs) in an interactive

process, in which the category variables for each class in the training set (*e.g.*, ±1) is used to optimise the model. A straight line that divides the classes' regions is then found [31].

Similarly to PLS, PCA also reduces the original data into a few set of variables called principal components (PCs). These variables are orthogonal to each other and account most of the explained variance from the original data set. They are composed of scores and loadings that are used to identify similarities/dissimilarities among the samples and the weight that each variable contributes for the PCA model, respectively [32]. However, differently from PLS, the category variables are not used for this reduction. To perform a supervised classification model, the PCA scores are employed as input variables for discriminant algorithms. This procedure avoids collinearity problems and also speeds up computational analysis.

LDA and QDA are discriminant algorithms that create a classification rule between the classes based on a Mahalanobis distance. The main difference between these techniques is that LDA uses a pooled covariance matrix to calculate the discriminant function between the classes, whereas QDA uses the variance-covariance matrix of each class separately [33]. Therefore, QDA usually achieves better performance than LDA when analysing complex data sets where the variance structures between the classes are very different. The LDA (L_{ik}) and QDA (Q_{ik}) classification scores are calculated following the equations [34]:

$$L_{ik} = (\mathbf{x}_i - \bar{\mathbf{x}}_k)^{\mathrm{T}} \sum_{\mathrm{pooled}}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_k) - 2\log_e \pi_k \quad (1)$$

$$Q_{ik} = (\mathbf{x}_i - \bar{\mathbf{x}}_k)^{\mathrm{T}} \sum_{k}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_k) + \log_e |\sum_k| - 2\log_e \pi_k \quad (2)$$

in which $\mathbf{x}_i$ is the vector containing the classification variables for sample i (*e.g.*, PCA scores for *A* components); $\bar{\mathbf{x}}_k$ is the mean vector of class *k*; $\sum_k$ is the variance-covariance matrix of class *k*; $\sum_{\mathrm{pooled}}$ is the pooled covariance matrix; and π_k is the prior probability of class *k*. These last three terms are calculated by [34]:

$$\sum_k = \frac{1}{n_k - 1} \sum_{i=1}^{n_k} (\mathbf{x}_i - \bar{\mathbf{x}}_k)\,(\mathbf{x}_i - \bar{\mathbf{x}}_k)^{\mathrm{T}} \quad (3)$$

$$\sum_{\text{pooled}} = \frac{1}{n} \sum_{k=1}^{K} n_k \sum_k \quad (4)$$

$$\pi_k = \frac{n_k}{n} \quad (5)$$

where n_k is the number of samples of class *k*; n is the total number of samples in the training set; and K is the number of classes.

On the other hand, SVM is a technique that classifies data sets in a completely non-linear fashion. For this, SMVs classifiers work by finding a classification hyperplane that separates the data clusters providing the largest margin of separation [35]. During model construction, the data is transformed into a different feature space by means of a kernel function that is responsible for the SVM classification ability [33]. The most common kernel function is the radial basis function (RBF). The SVM classifier takes the form of [36]:

$$f(x) = \text{sign}\left(\sum_{i=1}^{N_{SV}} \alpha_i y_i K(\mathbf{x}_i, \mathbf{z}_j) + b\right) \quad (6)$$

where N_{SV} is the number of support vectors; α_i is the Lagrange multiplier; y_i is the class membership (*e.g.*, ±1); b is the bias parameter; and $K(\mathbf{x}_i, \mathbf{z}_j)$ is the RBF kernel function, calculated by:

$$K(\mathbf{x}_i, \mathbf{z}_j) = \exp\left(-\gamma \left\|\mathbf{x}_i - \mathbf{z}_j\right\|^2\right) \quad (7)$$

in which $\mathbf{x}_i$ and $\mathbf{z}_j$ are samples measurement vectors; and γ is the parameter that determines the RBF width.

Results and Discussion

By employing the above-mentioned multivariate techniques (PCA-LDA, PLS-DA, PCA-QDA and PCA-SVM), it was demonstrated that some provided superior performance than others.

The techniques were very different from each other and were used following a parsimonious order (PCA-LDA < PLS-DA < PCA-QDA < PCA-SVM). It is natural to expect an improvement of the results when more robust algorithms are applied, as the classification methods varied from a linear (PCA-LDA and PLS-DA) to a completely non-linear classification algorithm (PCA-SVM). Analysis of the plasma samples deposited on Al foil showed classification to be: 68% sensitivity and 70% specificity (68% accuracy) after PLS-DA; 47% sensitivity and 75% specificity after PCA-LDA (51% accuracy); 83% sensitivity and 45% specificity after PCA-QDA (78% accuracy); 100% sensitivity and 90% specificity (98% accuracy) after PCA-SVM. For plasma samples that were deposited on low-E slides the results were: 65% sensitivity and 65% specificity (65% accuracy) after PLS-DA; 46% sensitivity and 85% specificity (52% accuracy) after PCA-LDA; 96% sensitivity and 15% specificity (84% accuracy) after PCA-QDA; 100% sensitivity and 85% specificity (98% accuracy) after PCA-SVM (Table 1). All PCA-based models for plasma samples were built with 10 PCs, accounting 99.11% and 96.84% of explained variance for Al and low-E substrates, respectively. PLS-DA was built with 10 LVs, accounting 98.97% and 95.28% of explained variance for Al and low-E substrates, respectively.

After applying classification algorithms for the blood serum samples, the results using Al foil as a substrate were: 82% sensitivity and 75% specificity (81% accuracy) after PLS-DA; 90% sensitivity and 40% specificity (81% accuracy) after PCA-LDA; 94% sensitivity and 50% specificity (86% accuracy) after PCA-QDA; 100% sensitivity and 70% specificity (94% accuracy) after PCA-SVM. When using serum samples on low-E slides the results were: 78% sensitivity and 90% specificity (80% accuracy) after PLS-DA; 63% sensitivity and 50% specificity (61% accuracy) after PCA-LDA; 97% sensitivity and 20% specificity (83% accuracy) after PCA-QDA; 100% sensitivity and 75% specificity (95% accuracy) after PCA-SVM (Table 2). All PCA-based models for serum samples were built with 10 PCs, accounting

for 98.78% and 97.50% of explained variance for Al and low-E substrates, respectively. PLS-DA was built with 10 LVs, accounting for 98.43% and 90.24% of explained variance for Al and low-E substrates, respectively.

Overall, PCA-SVM was found to provide optimal results for both plasma and serum samples regardless of the substrate that was used (Fig. 1 and 2). This was due to the fact that PCA-SVM can create a more complex decision boundary between the classes, classifying even non-linearly separable data [33, 35]. In addition, SVM creates large margins of separation between the classes, which provides more stability to the classifier. In this sense, small disturbances or noises do not cause misclassification [35]. Standard deviation (SD) was higher for Al foil in comparison to low-E slides (Fig. 1 and 2). This improved the Al foil classification models as more sources of variation were contemplated during model construction, thus creating well-distributed boundary functions and increasing the robustness of the classification. The SD in the training set decreases the degree of overfitting and provides better predictive capacity [37]. The PCA-SVM cost function and optimization to estimate RBF parameters are shown in Fig. 3, where the red 'X' mark represents the optimal value. This optimization was performed in order to avoid overfitting and to ensure classification stability. Fig. 4 shows the reference and predicted class labels (1 for control; and 2 for cancer) using PCA-SVM with the samples from the test set; if the yellow (predicted) and blue (reference) lines are superposed, then the values are equal (*i.e.*, no misclassification). For all substrates and type of samples (plasma and serum), there was no misclassification in the cancer set, reflecting the 100% sensitivity of PCA-SVM models. A degree of misclassification was observed in the control set, particularly when using serum samples. More specifically, specificity was higher in Al foil (90%) in contrast to low-E (85%); this has provided the slightly higher accuracy in Al foil (98.5%) in contrast to low-E (97.7%), in the plasma dataset. This can be seen in Fig. 4A and 4B as there are two and three misclassified spectra, respectively ("continuous" peaks represent more than one spectrum). The

specificity differed slightly in the serum dataset too, when Al foil (70%) and low-E (75%) were used. This contributed to the slightly lower accuracy in Al foil (94.5%) in contrast to low-E (95.5%). In this case, there were six misclassified spectra for Al foil and five for low-E slides. Although both PCA-LDA and PCA-QDA were regularized to correct classes having different sizes (prior probability term in eq. 1 and 2), the number of errors is larger on the smaller class (healthy control) due to the influence of the unequal class sizes to the classifiers. To summarise, Al foil has been seen to perform better than low-E in the plasma dataset, while in the serum dataset it achieved slightly lower specificity, but still high enough and comparable to low-E.

PCA-SVM models (Fig. 5) have different loadings profiles according to the type of sample and substrate. The loadings are dependent on the nature of the dataset used for the PCA model and they can differ depending on the input. Even though the same sample type is used, the change of the substrate has subsequently changed the spectral profile as well. Any variation above the instrumental noise can cause variation in the loading profiles. For instance, differentiation was also observed at specific spectral peaks between Al foil and low-E substrates (Fig. S1). Even though some spectral regions were visually similar, the reasoning of using multivariate analysis is to overcome visual interpretation which can be inaccurate. Therefore, a statistical *t*-test (95% confidence level) has been performed to calculate p-values for each spectral point between Al foil and low-E as well as between plasma and serum. The results showed that many wavenumbers were statistically significant ($p < 0.05$, 95% confidence level) irrespectively of the visual similarities (Fig. S2). Additionally, the fact that PC1 accounted for low values of explained variance (70.09% for plasma-Al; 38.98% for plasma low-E; 69.48% for serum-Al; and 28.69% for serum low-E) due to the high complexity of the biological dataset, makes the loadings interpretation even harder.

Using aluminium as substrate, larger coefficients were found between ~1000-1150 cm^{-1} for both plasma and serum samples, indicating possible glycogen and phosphate absorptions;

between ~1400-1480 cm^{-1}, indicating possible stretching vibrations of COO^- groups in fatty acids and amino acids; at ~1504 cm^{-1} for serum, signalling amide II absorption; and at ~1744 cm^{-1} for plasma, indicating C=O stretching of lipids [38]. Using low-E slides as substrate for plasma samples, larger coefficients were found at ~1628 cm^{-1} (amide I), ~1655 cm^{-1} (amide I) and ~1744 cm^{-1} (C=O stretching of lipids); whereas for serum samples, the coefficients were greater at ~1504 cm^{-1} (amide II), ~1620 cm^{-1} (base carbonyl stretching and ring breathing mode in nucleic acids) and 1655 cm^{-1} (amide I) [38]. Such absorptions are known for signalling biological changes using mid-IR spectroscopy [19].

The classification accuracies achieved for the segregation between endometrial cancer patients and controls are remarkably high (~95-98%), suggesting that blood-based ATR-FTIR spectroscopy could potentially be an accurate and objective diagnostic tool for endometrial cancer. Investigation of a panel of multiple biomolecules could be the reason for the achieved accuracies. Several molecular biomarkers have been suggested over the years, such as carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), cancer antigen 15-3 (CA15-3), immunosuppressive acidic protein (IAP), human epididymis protein-4 (HE4), apolipoprotein-1 (ApoA-1), prealbumin (TTR) and transferrin (TF); a combination of CA125 and HE4 has also been implied to improve diagnosis and classification of the disease [39-42]. However, the resulting sensitivities and specificities of the above-mentioned biomarkers are low, rendering them clinically unusable. Therefore, spectroscopic methods are ideal, as they can simultaneously extract information from a range of molecules. Another possible rationale behind the diagnostic results could be the existence of circulating tumour DNA (ctDNA) fragments in the bloodstream of cancer patients, which would make them distinct from the normal population [43, 44]. Nowadays, ctDNA is increasingly investigated and is considered to be useful as a biomarker for malignancy cases [45]. Nevertheless, for an accurate and specific biomarker detection, vibrational spectroscopy would need to be complemented with other

techniques as well, or maybe make use of labels or antibodies that would be molecule specific. IR spectroscopy alone indicates some molecular fragments which are indicative of biomolecules, such as proteins, lipids or carbohydrates. However, each spectral peak may 'hide' more than one molecules and thus, it is not preferred to assign specific biomarkers to specific peaks.

In this study, plasma samples resulted in slightly higher diagnostic accuracies (~98%) in contrast to serum samples (~95%). Current studies are unclear on whether serum or plasma is a better source for ctDNA [44]. However, plasma has been previously found superior and the specificity obtained using serum has been related to a higher concentration of normal cell-free DNA (cfDNA), produced by the lysis of white blood cells during clotting [46, 47]. This could potentially justify the lower classification rates found when using serum.

Careful consideration of the substrate, onto which the biological sample is placed, is critical in order to collect reproducible and high-quality spectra. When comparing the classification results coming from Al foil and low-E slides (Fig. 1 and 2), it is clear that Al foil not only achieved equally high results with low-E but, in the plasma dataset, it even provided slightly higher sensitivities and specificities (Fig. 1). Previous work has indicated that Al foil generates no background noise, leaving the quality of the biological spectra unaffected; our study used a larger number of subjects, which was needed to verify these preliminary results and also study the impact on the sensitivity and specificity. Studies have also demonstrated the enhancement of the IR signal in ATR mode when the sample is deposited onto metal surfaces creating a similar effect to surface enhanced Raman spectroscopy (SERS), which has been given the name surface enhanced IR absorption spectroscopy (SEIRAS). Molecules on metal surfaces show 10-100 times stronger signal that without the metal [48-51] and on the basis of this we have hypothesized that Al foil slides may also promote this effect. However, this requires further and more detailed investigation that will be the focus of a future study. The economic

cost of low-E slides has been estimated before and is not extremely high, especially when compared with substrates like CaF_2 and Au-coated slides [21]. Nonetheless, when it becomes a matter of routine use, in a clinical setting for instance, the annual cost becomes considerably high and this could render biospectroscopy prohibitive for translation into clinical practice. The fact that Al foil slides are suitable for both IR and Raman studies is also an important advantage as it would ease clinical implementation. The results of our study have shown that Al foil slides could make an ideal, cost-effective substrate for biomedical studies employing vibrational spectroscopy.

Conclusion

To summarise, biospectroscopy could potentially be used as a screening tool for endometrial cancer in postmenopausal women as it provides exceptionally high sensitivities and specificities with a simple blood test. This could automatically enable a large number of women to be assessed on a daily basis. Using disposable, low-cost and, at the same time, high-performance substrates would allow for universal studies with thousands of participants; this would probably also generate an interest for multi-centre studies which could further validate the pre-analytical, analytical and post-analytical phases of biospectroscopy.

Acknowledgements

MP would like to thank Rosemere Cancer Foundation for funding and Giouli Paraskevaidou for providing effective feedback. CLMM would like to acknowledge CAPES/Doutorado Pleno no Exterior/No. 88881.128982/2016-01 for financial support. KMGL acknowledges CNPq (grant 305962/2014-0) for financial support.

Figures

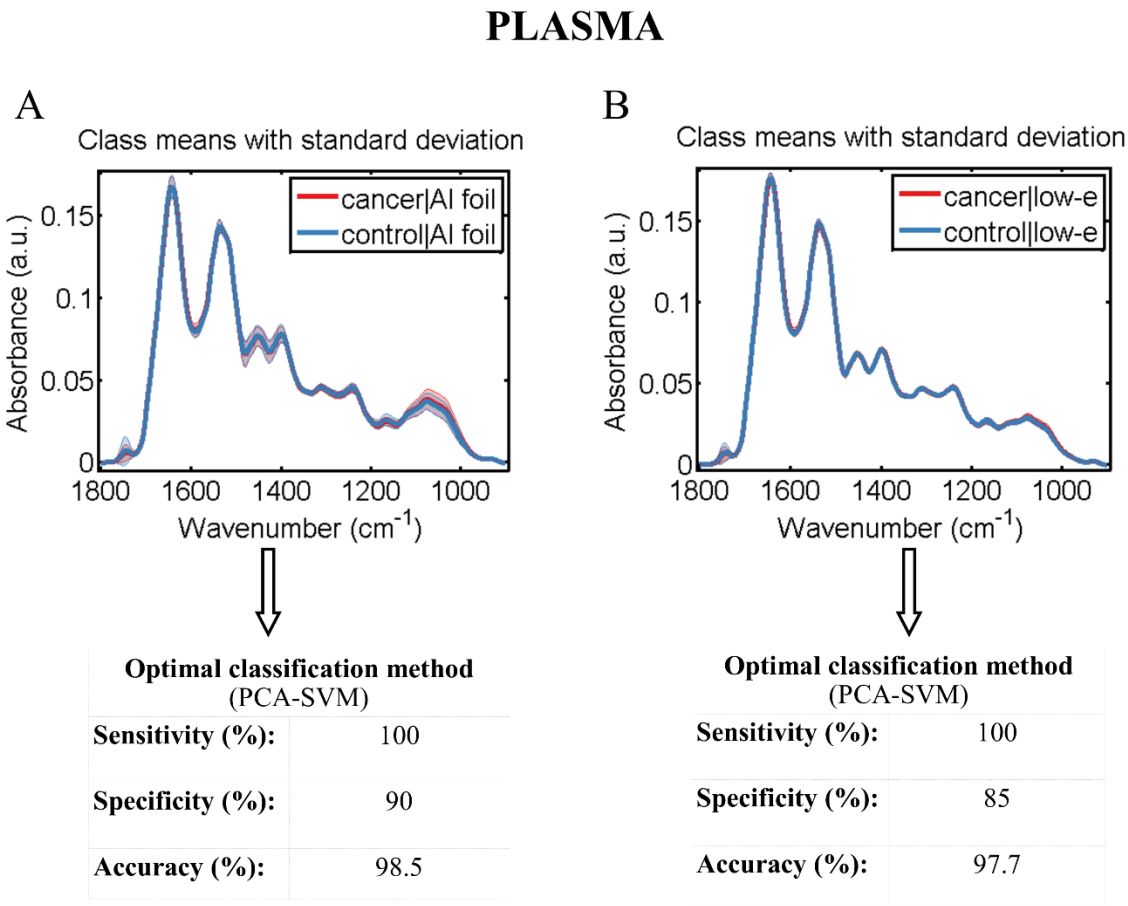

Figure 1: Pre-processed spectra of plasma comparing endometrial cancer (n=70) with controls (n=15). (A) Endometrial cancer versus healthy controls; samples were analysed on aluminium (Al) foil. Sensitivity and specificity were 100% and 90%, respectively. (B) Endometrial cancer versus healthy controls; samples analysed on low-E slides. Sensitivity and specificity were 100% and 85%, respectively.

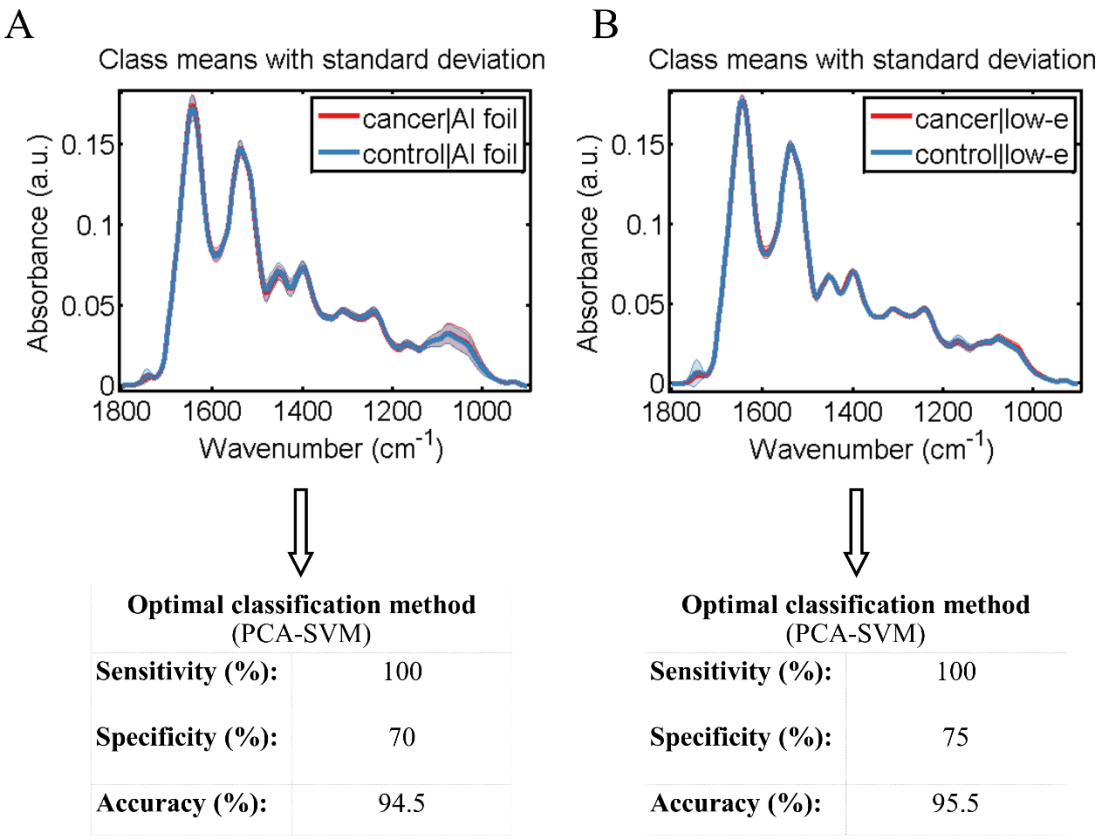

Figure 2: Pre-processed spectra of serum comparing endometrial cancer (n=60) with controls (n=15). (A) Endometrial cancer versus healthy controls; samples were analysed on aluminium (Al) foil. Sensitivity and specificity were 100% and 70%, respectively. (B) Endometrial cancer versus healthy controls; samples analysed on low-E slides. Sensitivity and specificity were 100% and 75%, respectively.

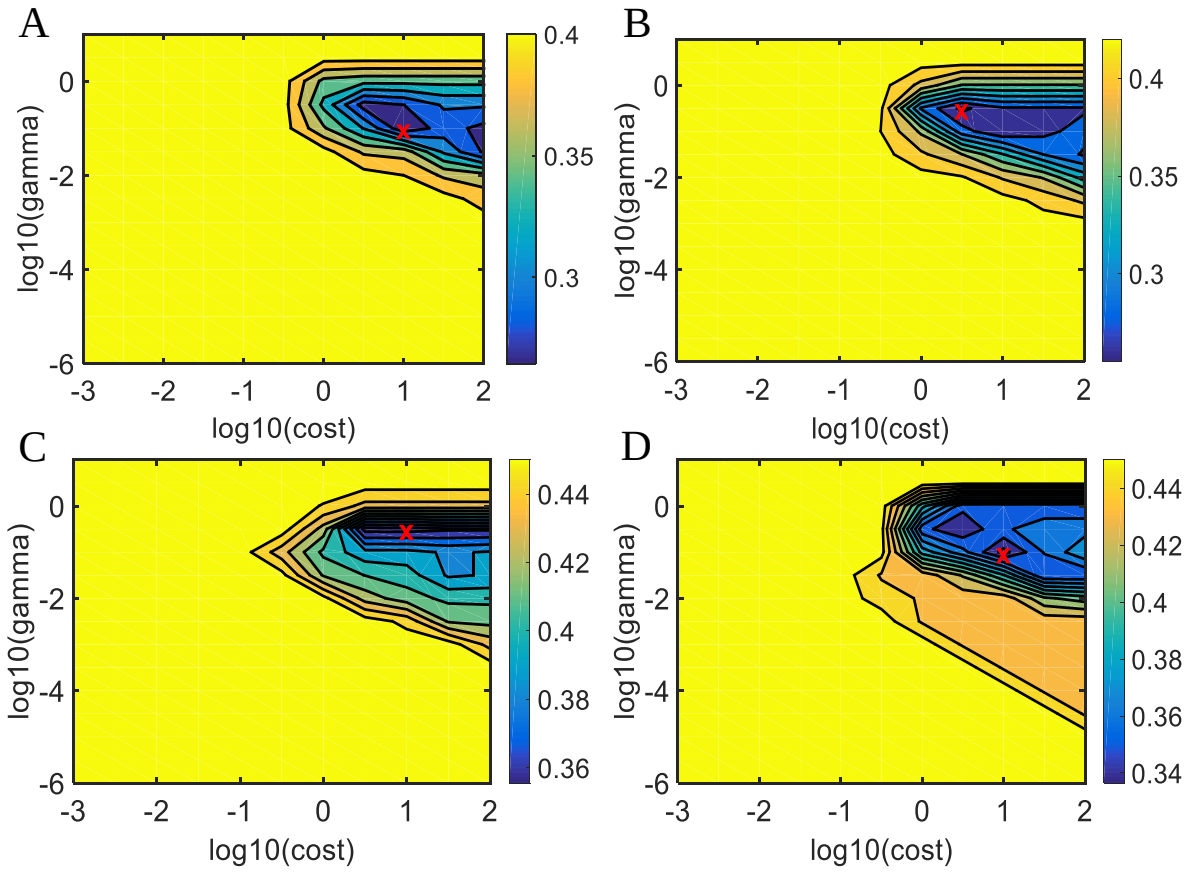

Figure 3: PCA-SVM cost function and radial basis function (RBF) parameter optimization. (A) Plasma samples with aluminium (Al) foil as a substrate. (B) Plasma samples with low-E slides as a substrate. (C) Serum samples with aluminium (Al) foil as substrate. (D) Serum samples with low-E slides as substrate. Gamma: RBF parameter (γ). Colour bar: misclassification rate using cross-validation.

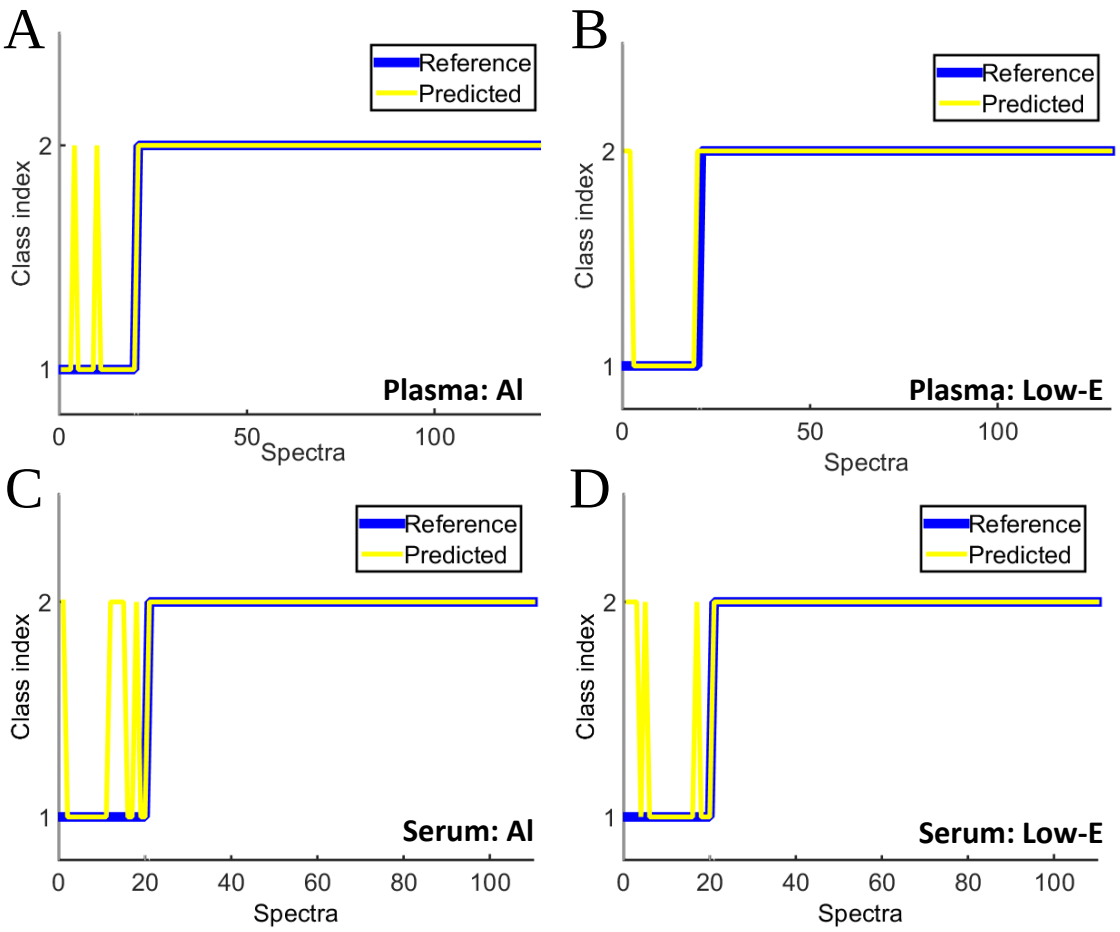

Figure 4: Reference and predicted class labels using PCA-SVM in the test set. (A) Plasma samples with aluminium (Al) foil as a substrate; sensitivity was 100% and specificity 90% (two misclassified spectra). (B) Plasma samples with low-E slides as a substrate; 100% sensitivity and 85% specificity (three misclassified spectra). (C) Serum samples with aluminium (Al) foil as substrate; 100% sensitivity and 70% specificity (six misclassified spectra). (D) Serum samples with low-E slides as substrate; 100% sensitivity and 75% specificity (five misclassified spectra). Class 1 = control; and class 2 = cancer.

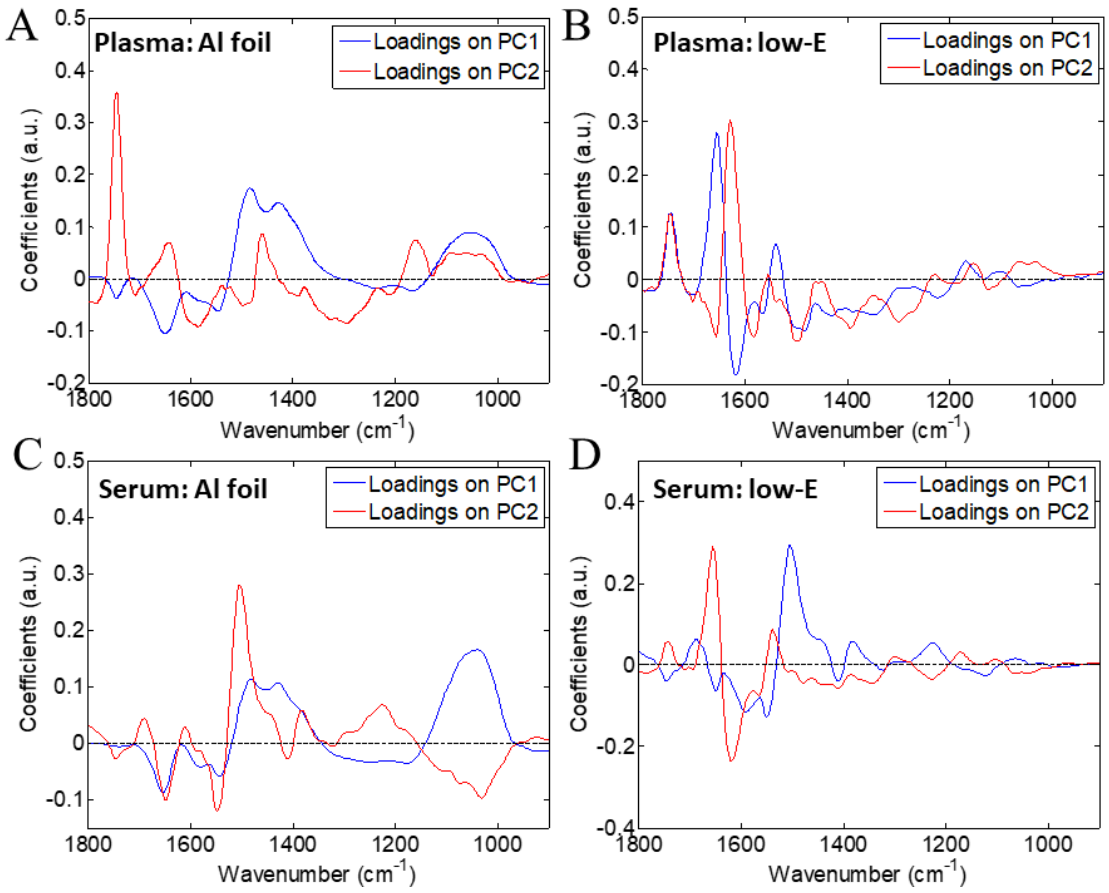

Figure 5: Loading plots generated after PCA analysis. (A) Loadings on PC1, PC2 for plasma samples deposited on aluminium (Al) foil slides. (B) Loadings on PC1, PC2 for plasma samples deposited on low-E slides. (C) Loadings on PC1, PC2 for serum samples deposited on aluminium (Al) foil slides. (D) Loadings on PC1, PC2 for serum samples deposited on low-E slides.

Tables

Table 1: Classification algorithms applied after the analysis of blood plasma samples. Results for both substrates, aluminium foil and low-E slide, are shown below.

Correct classification rate (%):

	Training (%)	Validation (%)	Test (%)
Aluminium foil			
PLS-DA	69.1	64.5	68.5
PCA-LDA	67.8	65.0	51.5
PCA-QDA	85.2	80.0	77.7
PCA-SVM	99.0	93.3	98.5
Low-E			
PLS-DA	71.1	71.8	65.4
PCA-LDA	62.7	54.2	52.3
PCA-QDA	85.2	82.5	83.8
PCA-SVM	99.8	97.5	97.7

Quality parameters (%):

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Aluminium foil			
PLS-DA	68.5	68.2	70.0
PCA-LDA	51.5	47.3	75.0
PCA-QDA	77.7	83.6	45.0
PCA-SVM	98.5	100	90.0
Low-E			
PLS-DA	65.4	65.5	65.0
PCA-LDA	52.3	46.4	85.0
PCA-QDA	83.8	96.4	15.0
PCA-SVM	97.7	100	85.0

Table 2: Classification algorithms applied after the analysis of blood serum samples. Results for both substrates, aluminium foil and low-E slide, are shown below.

SERUM

Correct classification rate (%):

	Training (%)	Validation (%)	Test (%)
Aluminium foil			
PLS-DA	80.0	79.1	80.9
PCA-LDA	72.1	79.1	80.9
PCA-QDA	84.3	79.1	86.4
PCA-SVM	98.3	93.6	94.5
Low-E			
PLS-DA	85.7	71.8	80.0
PCA-LDA	70.2	65.5	60.9
PCA-QDA	84.2	88.2	82.7
PCA-SVM	99.1	98.2	95.5

Quality parameters (%):

	Accuracy (%)	Sensitivity (%)	Specificity (%)
Aluminium foil			
PLS-DA	80.9	82.2	75.0
PCA-LDA	80.9	90.0	40.0
PCA-QDA	86.4	94.4	50.0
PCA-SVM	94.5	100	70.0
Low-E			
PLS-DA	80.0	77.8	90.0
PCA-LDA	60.9	63.3	50.0
PCA-QDA	82.7	96.7	20.0
PCA-SVM	95.5	100	75.0

References

[1] G. Graça, A. S. Moreira, A. J. V. Correia, B. J. Goodfellow, A. S. Barros, I. F. Duarte, I. M. Carreira, E. Galhano, C. Pita, M. d. C. Almeida, A. M. Gil *Anal Chim Acta*. **2013**, *764*, 24-31.

[2] A. Khoshmanesh, M. W. A. Dixon, S. Kenny, L. Tilley, D. McNaughton, B. R. Wood *Anal Chem*. **2014**, *86*, 4379-4386.

[3] S. E. Taylor, K. T. Cheung, I. I. Patel, J. Trevisan, H. F. Stringfellow, K. M. Ashton, N. J. Wood, P. J. Keating, P. L. Martin-Hirsch, F. L. Martin *Br J Cancer*. **2011**, *104*, 790-797.

[4] E. Staniszewska, K. Malek, M. Baranska *Spectrochim Acta Mol Biomol Spectrosc*. **2014**, *118*, 981-986.

[5] K. Gajjar, L. D. Heppenstall, W. Pang, K. M. Ashton, J. Trevisan, I. I. Patel, V. Llabjani, H. F. Stringfellow, P. L. Martin-Hirsch, T. Dawson, F. L. Martin *Anal Methods*. **2013**, *5*, 89-102.

[6] N. Stone, R. Baker, K. Rogers, A. W. Parker, P. Matousek *Analyst*. **2007**, *132*, 899-905.

[7] F. M. Lyng, E. Ó. Faoláin, J. Conroy, A. Meade, P. Knief, B. Duffy, M. Hunter, J. Byrne, P. Kelehan, H. Byrne *Exp Mol Pathol*. **2007**, *82*, 121-129.

[8] J. R. Hands, G. Clemens, R. Stables, K. Ashton, A. Brodbelt, C. Davis, T. P. Dawson, M. D. Jenkinson, R. W. Lea, C. Walker, M. J. Baker *J Neurooncol*. **2016**, *127*, 463-472.

[9] P. Carmona, M. Molina, M. Calero, F. Bermejo-Pareja, P. Martinez-Martin, A. Toledano *J Alzheimers Dis*. **2013**, *34*, 911-920.

[10] M. Paraskevaidi, C. L. M. Morais, K. M. G. Lima, J. S. Snowden, J. A. Saxon, A. M. T. Richardson, M. Jones, D. M. A. Mann, D. Allsop, P. L. Martin-Hirsch, F. L. Martin *Proc Natl Acad Sci USA*. **2017**, *114*, E7929-e7938.

[11] A. Shapiro, O. N. Gofrit, G. Pizov, J. K. Cohen, J. Maier *Eur Urol*. **2011**, *59*, 106-112.

[12] K. Maheedhar, R. A. Bhat, R. Malini, N. B. Prathima, P. Keerthi, P. Kushtagi, C. M. Krishna *Photomed Laser Surg*. **2008**, *26*, 83-90.

[13] K. Guze, H. C. Pawluk, M. Short, H. Zeng, J. Lorch, C. Norris, S. Sonis *Head Neck*. **2015**, *37*, 511-517.

[14] A. Sahu, K. Dalal, S. Naglot, P. Aggarwal, C. M. Krishna *Plos One*. **2013**, *8*, e78921.

[15] J. Zheng, L. He *Compr Rev Food Sci Food Saf*. **2014**, *13*, 317-328.

[16] L. T. Kerr, H. J. Byrne, B. M. Hennelly *Anal Methods*. **2015**, *7*, 5041-5052.

[17] B. D. Beier, A. J. Berger *Analyst*. **2009**, *134*, 1198-1202.

[18] H. J. Byrne, P. Knief, M. E. Keating, F. Bonnier *Chem Soc Rev*. **2016**, *45*, 1865-1878.

[19] M. J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, C. Hughes, P. Lasch, P. L. Martin-Hirsch, B. Obinaju, G. D. Sockalingum, J. Sulé-Suso, R. J. Strong, M. J. Walsh, B. R. Wood, P. Gardner, F. L. Martin *Nat Protoc*. **2014**, *9*, 1771-1791.

[20] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone, F. L. Martin *Nat. Protoc*. **2016**, *11*, 664-687.

[21] L. Cui, H. J. Butler, P. L. Martin-Hirsch, F. L. Martin *Anal Methods*. **2016**, *8*, 481-487.

[22] M. Paraskevaidi, P. L. Martin-Hirsch and F. L. Martin, in: *Characterization Tools for Nanoscience and Nanomaterials*, Springer, **2017** (in press).

[23] K. Gajjar, J. Trevisan, G. Owens, P. J. Keating, N. J. Wood, H. F. Stringfellow, P. L. Martin-Hirsch, F. L. Martin *Analyst*. **2013**, *138*, 3917-3926.

[24] K. Wehbe, J. Filik, M. D. Frogley, G. Cinque *Anal Bioanal Chem.* **2013**, *405*, 1311-1324.

[25] K. A. Heys, R. F. Shore, M. G. Pereira, F. L. Martin *Environ Sci Technol*. **2017**, *51*, 8672-8681.

[26] Cancer Research UK, http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/uterine-cancer/incidence#heading-Ten, Accessed February 2018.

[27] M. J. Baker, S. R. Hussain, L. Lovergne, V. Untereiner, C. Hughes, R. A. Lukaszewski, G. Thiefin, G. D. Sockalingum *Chem Soc Rev*. **2016**, *45*, 1803-1818.

[28] J. Trevisan, P. P. Angelov, P. L. Carmichael, A. D. Scott, F. L. Martin *Analyst*. **2012**, *137*, 3202-3215.
[29] R. W. Kennard, L. A. Stone *Technometrics*. **1969**, *11*, 137-148.
[30] D. B. Hibbert *Pure Appl Chem*. **2016**, *88*, 407-443.
[31] R. G. Brereton, G. R. Lloyd *J Chemometrics*. **2014**, *28*, 213-225.
[32] R. Bro, A. K. Smilde *Anal Methods*. **2014**, *6*, 2812-2831.
[33] S. J. Dixon, R. G. Brereton *Chemom Intell Lab Syst*. **2009**, *95*, 1-17.
[34] F. S. Costa, P. Silva, C. Lelis, R. Theodoro, T. Arantes, K. de Lima *Anal Methods*. **2017**.
[35] P. de Boves Harrington *Anal Chem*. **2015**, *87*, 11065-11071.
[36] C. L. Morais, F. S. Costa, K. M. Lima *Anal Methods*. **2017**, *9*, 2964-2970.
[37] H. Martens, P. Geladi, Multivariate calibration, Wiley Online Library, **1989**.
[38] Z. Movasaghi, S. Rehman, I. U. Rehman *Appl Spectrosc Rev*. **2007**, *42*, 493-541.
[39] L. Zanotti, E. Bignotti, S. Calza, E. Bandiera, G. Ruggeri, C. Galli, G. Tognon, M. Ragnoli, C. Romani, R. A. Tassi *Clin Chem Lab Med*. **2012**, *50*, 2189-2198.
[40] P. B. Panici, G. Scambia, G. Baiocchi, L. Perrone, S. Greggi, F. Battaglia, S. Mancuso *Gynecol Obstet Invest*. **1989**, *27*, 208-212.
[41] Y. Ueda, T. Enomoto, T. Kimura, T. Miyatake, K. Yoshino, M. Fujita, T. Kimura *Cancers*. **2010**, *2*, 1312.
[42] G. Farias-Eisner, F. Su, T. Robbins, J. Kotlerman, S. Reddy, R. Farias-Eisner *Am J Obstet Gynecol*. **2010**, *202*, 73.e71-75.
[43] C. Bettegowda, M. Sausen, R. J. Leary, I. Kinde, Y. Wang, N. Agrawal, B. R. Bartlett, H. Wang, B. Luber, R. M. Alani *Sci Transl Med*. **2014**, *6*, 224ra224-224ra224.
[44] Y. I. Elshimali, H. Khaddour, M. Sarkissyan, Y. Wu, J. V. Vadgama *Int J Mol Sci*. **2013**, *14*, 18925-18958.
[45] L. A. Diaz Jr, A. Bardelli *J Clin Oncol.* **2014**, *32*, 579-586.
[46] A. Vallée, M. Marcq, A. Bizieux, C. E. Kouri, H. Lacroix, J. Bennouna, J.-Y. Douillard, M. G. Denis *Lung Cancer*. **2013**, *82*, 373-374.
[47] S. El Messaoudi, F. Rolet, F. Mouliere, A. R. Thierry *Clinica Chimica Acta*. **2013**, *424*, 222-230.
[48] S. E. Glassford, B. Byrne, S. G. Kazarian *Biochim Biophys Acta - Proteins Proteom*. **2013**, *1834*, 2849-2858.
[49] K. Ataka, S. T. Stripp, J. Heberle *Biochim Biophys Acta – Biomembranes*. **2013**, *1828*, 2283-2293.
[50] J.-Y. Xu, T.-W. Chen, W.-J. Bao, K. Wang, X.-H. Xia *Langmuir*. **2012**, *28*, 17564-17570.
[51] R. Adato, H. Altug *Nat Commun.* **2013**, *4*, 2154.