

Analyst

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

1
2
3 **Wavenumber selection based analysis in Raman spectroscopy improves skin cancer**
4
5 **diagnostic specificity†**
6
7
8
9

10 Jianhua Zhao,^{1,2} Haishan Zeng,^{1,2} Sunil Kalia¹ and Harvey Lui^{1,2}
11
12
13
14
15
16

17 ¹Photomedicine Institute, Department of Dermatology and Skin Science, University of British
18 Columbia and Vancouver Coastal Health Research Institute, Vancouver, Canada
19

20 ²Imaging Unit - Integrative Oncology Department, BC Cancer Agency Research Center,
21 Vancouver, Canada
22
23
24
25
26
27
28
29
30
31
32
33

34 **Corresponding author**
35

36 Haishan Zeng
37

38 Professor
39

40 Integrative Oncology Department – Imaging Unit
41

42 BC Cancer Agency Research Center
43

44 675 West 10th Ave
45
46
47

48 Vancouver, BC V5Z 1L3
49

50 Phone: (604)-675-8083
51

52 Email: hzeng@bccrc.ca
53
54
55
56
57
58
59
60

Abstract

Real-time Raman spectroscopy can be used to assist in assessing skin lesions suspicious for cancer. Most of the diagnostic algorithms are based on full band of the Raman spectra, either in the fingerprint region or the high wavenumber region. In this paper we explored wavenumber selection based analysis in Raman spectroscopy for skin cancer diagnosis. Wavenumber selection was implemented using windows of wavenumber and leave-one-out cross-validated stepwise regression or least and shrinkage selection operator (LASSO). The diagnostic algorithms were then generated from the selected windows of wavenumber using multivariate statistical analyses, including principal component and general discriminate analysis (PC-GDA) and partial least squares (PLS). In total a combined cohort of 645 confirmed lesions from 573 patients encompassing skin cancers, precancers and benign skin lesions were included, which were divided into training cohort (n = 518) and testing cohort (n = 127) according to the measurement time. It was found that the area under the receiver operating characteristic curve (ROC) was improved from 0.861 – 0.891 to 0.891 – 0.911 and the diagnostic specificity for fixed sensitivity 0.99-0.90 was improved from 0.17 – 0.65 to 0.20 – 0.75 with wavenumber selection based analysis.

1. Introduction

Raman spectroscopy is a noninvasive optical technique that can measure the vibrational modes of biomolecules, and has been investigated as a diagnostic tool for cancers of the skin¹⁻⁷, lung⁸⁻¹⁰, colon¹¹⁻¹³, stomach^{14, 15}, cervix¹⁶⁻²⁰, oral cavity^{21, 22}, and breast²³⁻²⁹. We recently reported three algorithms for various levels of skin cancer diagnosis based on a cohort of 518 cases of skin cancers and benign skin lesions¹. The areas of the receiver operating characteristic curve (ROC AUC) served as measures of diagnostic performance and were found to be in the range of 82-90%, corresponding to sensitivities of 90-99% and specificities of 15-68% depending on the diagnostic algorithm. These results were equivalent to or better than other photonics-based diagnostic technologies¹. A follow-up study involving 127 additional independent skin lesions validated our previous findings and confirmed that Raman spectroscopy was a reliable technique for *in vivo* skin cancer diagnosis². Our previous diagnostic algorithms were based on multivariate statistical analysis of the full-band of Raman spectra (including all the data points at each recorded wavenumber of the Raman spectra, i.e. 619 spectral measurements between 500 – 1800 cm⁻¹ for our system). In this paper we explore the feasibility of spectral analysis using selected wavenumbers versus full spectral data for improving Raman spectroscopy for skin cancer diagnosis.

Wavelength selection (hereafter interchangeably called wavenumber selection in Raman spectroscopy, is more commonly called variable selection or feature selection in Chemometrics) has been used for improving model predictions in agriculture and food science³⁰⁻³³. For example, Hoskuldsson et al presented a few strategies for variable and subset selection in partial least

1
2
3 squares (PLS) regression, including situations of high and low correlation between wavelength
4 and response variables³⁰. Norgaard et al proposed forward, backward and bi-direction interval
5 PLS models for regression³¹. Anderssen et al proposed a cross-model validation procedure to
6 reduce over optimism in variable selection³². Andersen et al summarized and compared the
7 performance of a number of variable selection techniques in regression including interval PLS,
8 genetic algorithms (GA), stepwise PLS, and LASSO (least and shrinkage selection operator)³³.
9 Bursac et al used a purposeful selection of variables in logistic regression, where the variables
10 were chosen from the most significant covariates first³⁴. All these algorithms have been used
11 successfully in regression-related problems.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 Some algorithms for wavelength selection have also been proposed for classification tasks³⁵⁻³⁸.
28 Most of these algorithms were nonlinear feature selection methods based on single wavelengths
29 and found some success in classifying Raman spectra³⁵⁻³⁸. Because the Raman signal at one
30 wavenumber (inverse of wavelength) is highly correlated to the signal at the neighboring
31 wavenumber, it is thought to be “wasteful and overly complex” to work on single wavenumbers
32 independently³³. Therefore, it is recommended to use windows of wavenumbers rather than a
33 single wavenumber³³. A number of studies selected wavenumbers according to fixed window
34 sizes. For example, Duraipandian et al reported that for the *in vivo* diagnosis of cervical
35 precancers using Raman spectroscopy GA-based wavenumber selection resulted in an
36 improvement of the area under the ROC curve (ROC AUC) from 81% to 85%²⁰. Li et al
37 reported nasopharyngeal cancer detection using Raman spectroscopy and GA-based
38 wavenumber selection and found that the ROC AUC was improved from 92% to 96%³⁹.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 detection and found the ROC AUC could be improved from 83% to 85-88% with wavenumber
4 selection⁴⁰. Nevertheless, the use of equal sized windows may not be optimal as the correlated
5 wavenumbers may be arbitrarily or unintentionally split into different windows. In this paper, we
6 propose a simple algorithm to implement windows of variable size, and explore wavenumber
7 selection based analysis in Raman spectroscopy for skin cancer diagnosis.
8
9
10
11
12
13
14
15
16
17
18

19 **2. Patients and methods**

20 **2.1. Patients and skin lesions**

21
22
23
24 The patient and lesion information in this study has been reported in previous publications where
25 the spectra were analyzed on a full band basis^{1,2}. In vivo Raman spectra of skin cancers and
26 benign skin lesions were acquired using the real-time Raman spectrometer with 1-second
27 integration time³. The raw Raman spectra from the spectrometer were then wavelength- and
28 intensity-calibrated³. Fluorescence background was removed using the fifth-order Vancouver
29 Raman Algorithm⁴¹. The background-corrected Raman spectra were normalized to their
30 respective area under the curve between 500 and 1800 cm^{-1} before statistical analysis. The
31 detailed distribution of lesions among clinical diagnostic subtypes is listed in **Table 1**. There are
32 two cohorts of patients: the first cohort consists of 518 sets of Raman spectra for benign and
33 malignant skin lesions from 453 patients for model training, which were acquired between
34 January 2003 and May 2011 (hereafter referred to as “cohort one” or “training cohort”); the
35 second cohort contains 127 sets of Raman spectra for separate benign and malignant skin lesions
36 from 120 patients for model testing, which were measured in the period between June 2011 and
37 May 2014 (hereafter referred to as “cohort two” or “testing cohort”). When the two cohorts of
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

patients are combined according to pathology, it is referred to as “combined cohort”. For the purpose of this study, each individual skin lesion of interest (represented by a single spectrum) was considered as an experimental unit for analysis.

2.2 Varying size of windows

A simple method was proposed to implement variably-sized windows. The size of windows is determined by the correlation coefficient between wavenumbers, given by,

$$r_{i,j} = \frac{\sum_{k=1}^N (R_{i,k} - \bar{R}_i)(R_{j,k} - \bar{R}_j)}{\sqrt{\sum_{k=1}^N (R_{i,k} - \bar{R}_i)^2 \sum_{k=1}^N (R_{j,k} - \bar{R}_j)^2}}$$

Where, $r_{i,j}$ is the correlation coefficient between wavenumber i and j , $R_{i,k}$ and $R_{j,k}$ is the Raman intensity at wavenumber i and j for lesion k , \bar{R}_i and \bar{R}_j are the mean Raman intensity at wavenumber i and j . N is the total number of lesions. A wavenumber and its neighboring wavenumber that have correlation coefficients greater than predetermined threshold can be classified into a single window. When the threshold is 1, the window size is reduced to 1 and the model is reduced to the original single wavenumber situation as previously reported.

2.3 Wavenumber selection

1
2
3 There are a number of wavenumber selection algorithms including stepwise regression, least
4 absolute shrinkage selection operator (LASSO) and genetic algorithms (GA)^{33, 42-45}. GA is a
5 method based on evolutionary theory and mimicks natural selection processes. The process is
6 completely random, and thus the final selected wavelengths are different for each simulation. It
7 has found some success in wavelength selection, but it bears higher risk for over fitting^{33, 42}. If
8 the final objective is classification, wavelength selection based on GA may not be appropriate. In
9 this paper, we will focus on the most classical stepwise regression and LASSO algorithm for
10 wavenumber selection based analysis⁴⁶. The wavenumber selection algorithm is a leave-one-out
11 cross-selection (LOO-CS) procedure: (1) one spectrum is left out, and the remaining spectra are
12 used for wavenumber selection; (2) a subset of the wavebands is identified that gives the best
13 classification or the smallest residual sum of squares; (3) the leave-one-out procedure is repeated
14 until all the spectra are tested; (4) the selected wavenumber is then accumulated for the above
15 leave-one-out procedure. The wavenumber that are selected more frequently during the above
16 leave-one-out procedure are used for final discrimination analysis.

37 38 39 **2.4 Classification**

40 Multi-variant statistical analysis methods including PC-GDA and PLS were applied to the
41 selected wavelengths for lesion classification⁴⁷. For a general PC-GDA analysis, the spectra are
42 divided into training spectra and testing spectra. The discrimination model is generated from the
43 training spectra, and is validated from the testing spectra independently. For a leave-one-out
44 cross-validated (LOO-CV) PC-GDA analysis, one spectrum is left out for testing, and the
45 remaining spectra are used for discrimination model training. The training spectra are scaled and
46 normalized by removing the mean and dividing by the standard deviation before applying the

1
2
3 PC-GDA analysis. The PC factors and the PC loadings of the training spectra are then calculated.
4
5 The PC factors derived from the training spectra are used to develop a general discrimination
6
7 model for classification where the classification of each lesion is known a priori. The testing
8
9 spectra are processed the same way as the training spectra for classification: the testing spectra
10
11 are scaled and normalized using the mean and standard deviation obtained from the training
12
13 spectra; the PC factors of the testing spectra are calculated based on the PC loadings of the
14
15 training spectra; and finally, the posterior probabilities of the testing spectra for skin cancer are
16
17 calculated based on the general discrimination model derived from the training spectra. The
18
19 above PC-GDA procedure is repeated until all the spectra are tested once. Lesions are then
20
21 classified based on the posterior probabilities. The procedures for PLS analysis are similar to the
22
23 above PC-GDA analysis.
24
25
26
27
28
29
30
31

32 **2.5 ROC curves**

33
34 ROC curves were calculated from the posterior probabilities derived from the above analysis and
35
36 represents the diagnostic performance of the model. The area under the ROC curve was a
37
38 measure of the diagnosis performance. The significance of these AUCs and comparisons
39
40 between different AUCs were also calculated according to Hanley and McNeil^{48, 49}. The p-value
41
42 between two ROC curves is calculated from z-score by^{48, 49}
43
44
45
46
47

$$48 \quad z = |A_1 - A_2| / \text{SQRT}(SE_1^2 + SE_2^2 - 2*\gamma*SE_1*SE_2).$$

49
50
51
52
53 where A_1 and A_2 are the areas under the two ROC curves, SE_1 and SE_2 are the standard errors of
54
55 the two ROC curves, and γ is the correlation between the two ROC curves, and SQRT is square
56
57
58
59
60

1
2
3 root function (Graphpad, La Jolla, California, USA) . For two unpaired datasets, $\gamma = 0$. [In
4
5 Microsoft Excel, a two-tail p-value is calculated as $p = 2*(1-NORMSDIST(z))$, where
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

root function (Graphpad, La Jolla, California, USA) . For two unpaired datasets, $\gamma = 0$. [In Microsoft Excel, a two-tail p-value is calculated as $p = 2*(1-NORMSDIST(z))$, where NORMSDIST is an excel built-in function]. All the ROC curves were calculated based on nonparametric techniques and were conducted separately for the PC-GDA and PLS analyses. For the purposes of comparing the different statistical methods as well as the performance of Raman spectroscopy versus other non-invasive diagnostic techniques, the specificities were calculated at fixed sensitivity levels at 90%, 95% and 99%. All the multivariate classification analyses in this study were implemented using MATLAB (version 2013b, Math-Works).

2.6 Cross validation

Two cross-validation tasks were implemented: (a) “cohort one” was used for wavenumber selection and algorithm generation in multivariate statistical analysis, while “cohort two” was used to testing the wavenumber selection algorithm and the discrimination models; (b) the “combined cohort” of all cases was used for wavenumber selection based on LOO-CS. The selected wavenumbers were then used for training and testing for classification based on LOO-CV. The former cross-validation task is a direct test of whether the wavenumber selection and discrimination algorithm derived from “cohort one” could be reliably applied to the independent “cohort two”. The latter cross-validation task is to test whether an increasing sample size (combined cohort) could improve wavenumber selection and diagnostic performance.

3. Results

3.1 Raman spectra

1
2
3 The averaged Raman spectra of skin cancers and precancers (including melanoma, basal cell
4 carcinoma, squamous cell carcinoma and actinic keratosis) and benign skin lesions (including
5 atypical nevus, blue nevus, compound nevus, intradermal nevus, junctional nevus and seborrheic
6 keratosis) of the combined cohort (n = 645) are shown in **Figure 1**. All the Raman spectra are
7 normalized to their respective areas under the curves between 500 cm^{-1} and 1800 cm^{-1} before
8 being averaged according to diagnosis. Similar to all our previous studies, the major Raman
9 bands for skin cancers and benign skin lesions are located around 855, 936, 1002, 1078, 1271,
10 1302, 1445 1655, and 1745 cm^{-1} . From the difference spectrum in **figure 1a** between skin
11 cancers and benign skin lesions, it can be seen that on average there are differences between skin
12 cancers and benign skin lesions at all the major Raman bands. Skin cancers show higher Raman
13 intensity at 1002, 1078, 1302, 1445 and 1655 cm^{-1} bands and lower intensity at 855, 936, 1271
14 and 1745 cm^{-1} bands based on the area-normalized Raman spectra. Major differences also occur
15 at trough regions such as $1330\text{-}1420\text{ cm}^{-1}$. These differences provide diagnostic information
16 between skin cancers and benign skin lesions.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 In order to determine the contribution of each waveband to the diagnosis, PLS analysis was
40 performed on the training cohort (n = 518) and the testing cohort (n = 127) for sequential,
41 nonoverlapping bands of 100 cm^{-1} width between $500\text{ - }1800\text{ cm}^{-1}$, and the results are shown in
42 **figure 1b**. The results for the training cohort were leave-one-out cross-validated. The results for
43 the testing cohort were based on the model generated in the independent training cohort. It was
44 found that the above Raman bands provided different diagnostic performance with the ROC
45 AUC ranging between 0.71 - 0.82 and 0.63 - 0.78 for the training cohort and the testing cohort
46 respectively. Selected Raman bands around $1200\text{ - }1300$, $1300\text{ - }1400\text{ cm}^{-1}$ provided better
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 diagnosis performance than other Raman bands. However, no single Raman band provided the
4
5 diagnostic performance as high as the full Raman band ($500 - 1800 \text{ cm}^{-1}$) where the ROC AUCs
6
7 were 0.896 and 0.889 for the training cohort and testing cohort respectively.
8
9

10 11 12 **3.2 Wavenumber selection by LASSO and classification by PC-GDA: cohort one for** 13 14 **training and cohort two for testing** 15 16

17 The correlation between wavenumber (Raman shift) of the training cohort of patients ($n = 518$)
18 was calculated. The diagnostic performance depends slightly on the correlation coefficients, with
19 the optimal correlation coefficient threshold to be 0.80-0.95 (**figure S1†**). When the correlation
20 coefficient threshold was set at 0.85, the Raman spectra of the training cohort were divided into
21 104 Raman bands of varying widths. The sizes of the windows ranged from 3 pixels to 27 pixels
22 ($5 - 45 \text{ cm}^{-1}$). After the Raman spectra were divided into 104 Raman bands, the training cohort
23 was used for wavenumber selection. Forty-seven Raman bands were selected by LASSO with
24 leave-one-out cross-validation, and are shown in **figure 1a** (circles on the difference spectrum
25 and the shaded grey bands). PC-GDA was then used for classification of the training cohort
26 based on leave-one-out cross-validation with the selected Raman bands. Once the classification
27 model was generated, it was applied to the testing cohort ($n = 127$) with the selected Raman
28 bands from the training cohort. **Figure 2a and 2b** showed the posterior probability for each
29 lesion to be classified as a skin cancer or precancers for the training cohort and the testing cohort
30 with the selected Raman bands, respectively. **Figure 2c** showed the ROC curves (solid line) with
31 95% confidence intervals (dashed lines) obtained from the distribution of posterior probabilities
32 in **figure 2a and 2b**. The ROC AUC for the training cohort with selected Raman bands was
33 found to be 0.905 (0.879 – 0.931). The ROC AUC for the testing cohort with selected Raman
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 bands based on the discrimination model generated from the training cohort was found to be
4
5 0.891 (0.835 – 0.948). The performance of the testing cohort was not statistically different from
6
7 the training cohort ($p = 0.6574$).
8
9

10
11
12 **Figures 2d-2f** are the respective posterior probability and ROC curves for the full Raman spectra
13 between 500-1800 cm^{-1} , which were reported in the previous studies ^{1,2}, and are presented here
14 for comparison purpose. **Figure 2d** showed the posterior probability for each lesion to be
15 classified by PC-GDA as a skin cancer or precancer for the training cohort based on LOOCV.
16
17 **Figure 2e** showed the posterior probability for each lesion to be classified as a skin cancer or
18 precancer for the testing cohort based on the discrimination model generated from the training
19 cohort. **Figure 2f** showed the ROC curves (solid line) with 95% CIs (dashed lines) obtained from
20 the distribution of posterior probabilities in **figure 2d-2e**. The ROC AUC for the training cohort
21 was 0.879 (0.829-0.929) and the ROC AUC for the testing cohort was 0.861 (0.796-0.927). The
22 difference in the ROC curves between the training cohort and the testing cohort were statistically
23 insignificant ($p = 0.6298$).
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 The ROC AUC for the training cohort was improved from 0.879 (0.829-0.929) to 0.905 (0.879 –
42 0.931) with wavenumber selection. The improvement was found to be statistically significant (p
43 < 0.0001). The ROC AUC for the testing cohort was improved from 0.861 (0.796-0.927) to
44 0.891 (0.835 – 0.948) with wavenumber selection. The improvement was also found to be
45 statistically significant ($p = 0.0057$). The diagnosis parameters including ROC AUCs and
46 sensitivities and specificities were summarized in **table 2** and **table 3**.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

3.3 Wavenumber selection by LASSO and classification by PC-GDA: combined cohort

The correlation between wavenumber (Raman shift) of the combined cohort of patients (n = 645) was calculated. The diagnostic performance depends slightly on the correlation coefficients, with the optimal correlation coefficient threshold to be 0.80-0.95. When the correlation coefficient threshold was set at 0.85, the Raman spectra of the combined cohort (n = 645) were also divided into 104 Raman bands of different sizes, similar to the results of the training cohort only. After the Raman spectra were divided into 104 Raman bands, the combined cohort was used for wavelength selection. Forty Raman bands were selected by LASSO with leave-one-out cross-validation. PC-GDA was then used for classification based on leave-one-out cross-validation with the selected Raman bands.

Figure 3 showed the wavelength selection and diagnostic results for the combined cohort (n = 645). **Figure 3a** and **3b** showed the posterior probability for each lesion to be classified as a skin cancer or precancer for the combined cohort without and with wavelength selection. **Figure 3c** showed the ROC curves (solid line) with 95% CIs (dashed lines) obtained from the distribution of posterior probabilities in **figure 3a and 3b**. The ROC AUC for the combined cohort with selected wavenumber was 0.906 (0.883-0.929), better than that with full wavenumber band where the ROC AUC was 0.891 (0.867-0.916). The improvement was statistically significant ($p < 0.0001$). The diagnosis parameters including ROC AUCs and sensitivities and specificities were also summarized in **table 2** and **table 3**.

3.4 Wavenumber selection by stepwise regression and classification by PC-GDA

1
2
3 Stepwise regression was also applied for wavenumber selection for the above two analyses.
4
5
6 When cohort one (n = 518) was used for wavenumber selection and training, and cohort two (n =
7
8 127) was used for testing, the area under the ROC curve was found to be 0.904 (0.878-0.930) for
9
10 the training cohort, and 0.911 (0.863-0.959) for the testing cohort. When the combined cohort (n
11
12 = 645) was analyzed, the ROC AUC was 0.899 (0.876 – 0.922) with wavenumber selection by
13
14 stepwise regression and classification by PC-GDA. The results are better than those of full
15
16 wavelength analysis where the ROC AUC was 0.891 (0.867-0.916). The improvement was
17
18 statistically significant (p = 0.0287). The diagnosis parameters including ROC AUCs and
19
20 sensitivities and specificities for stepwise regression and classification based on PC-GDA were
21
22 also summarized in **table 2** and **table 3**.
23
24
25
26
27
28

29 **3.5 Wavenumber selection with stepwise regression and LASSO, and classification with** 30 31 **PLS**

32
33
34 The above analyses after wavenumber selection by either stepwise regression or LASSO were
35
36 repeated with classification by PLS. The results were also summarized in **table 2 and 3**. The
37
38 results with PLS analyses are similar to those of PC-GDA analyses.
39
40
41
42

43 **3.6 Wavenumber selection based analysis improves diagnostic specificity**

44
45
46 **Figure 4a-4c** showed the diagnostic specificities of the above analyses at sensitivity levels of
47
48 0.99, 0.95 and 0.90 based on PC-GDA analysis, respectively. It could be seen that at 0.99 (0.96 –
49
50 1.00) sensitivity level, the specificity was improved from 0.17 (0.13 – 0.21) to 0.27 (0.22 – 0.33)
51
52 and 0.33 (0.28 – 0.39) for the training cohort, from 0.24 (0.13 – 0.38) to 0.20 (0.10 – 0.34) and
53
54 0.34 (0.21 – 0.49) for the testing cohort, and from 0.24 (0.20 – 0.29) to 0.26 (0.21 – 0.31) and
55
56
57
58
59
60

1
2
3 0.35 (0.26 – 0.40) for the combined cohort for LASSO-based and stepwise-based wavenumber
4 analysis, respectively. At 0.95 (0.91 – 0.97) sensitivity level, the specificity was improved from
5
6 0.41 (0.35 – 0.48) to 0.55 (0.49 – 0.61) and 0.56 (0.50 – 0.61) for the training cohort, from 0.48
7
8 (0.34 – 0.63) to 0.48 (0.34 – 0.63) and 0.54 (0.39 – 0.68) for the testing cohort, and from 0.54
9
10 (0.48 – 0.59) to 0.55 (0.50 – 0.61) and 0.54 (0.48 – 0.59) for the combined cohort for LASSO-
11
12 based and stepwise-based wavenumber analysis, respectively. At 0.90 (0.85 – 0.93) sensitivity
13
14 level, the specificity was improved from 0.64 (0.58 – 0.70) to 0.72 (0.67 – 0.77) and 0.75 (0.70 –
15
16 0.80) for the training cohort, from 0.54 (0.39 – 0.68) to 0.74 (0.60 – 0.85) and 0.64 (0.49 – 0.77)
17
18 for the testing cohort, and from 0.65 (0.60 – 0.70) to 0.71 (0.65 – 0.75) and 0.71 (0.66 – 0.76) for
19
20 the combined cohort for LASSO-based and stepwise-based wavenumber analysis, respectively.
21
22 Overall, the specificity was improved from 0.17 – 0.65 to 0.20 – 0.75 with wavenumber selection
23
24 based analysis. Diagnosis based on PLS generated similar results as the above PC-GDA analysis
25
26 (table 3).
27
28
29
30
31
32
33
34
35
36

37 4. Discussion

38 Raman spectroscopy has been used for cancer diagnosis including skin in a number of studies.
39
40 Most of the prior diagnoses were based on the full wavelength range of Raman spectra either in
41
42 the fingerprint region or the high wavenumber region¹⁻⁷. In an earlier publication we explored
43
44 real-time Raman spectroscopy for *in vivo* skin cancer diagnosis using the most conservative
45
46 multivariate statistical methods with the full band of spectra in the fingerprint region¹. It
47
48 established that the diagnostic performance of real-time Raman spectroscopy is equivalent to or
49
50 better than other diagnostic aids. A following independent study validated all the previous
51
52
53
54
55
56
57
58
59
60

1
2
3 findings ². However, the diagnostic specificity is still limited at high sensitivity levels. From the
4
5 clinical diagnostic point of view, it is very important to improve specificity at high sensitivity
6
7 levels to reduce unnecessary biopsies.
8
9

10
11
12 In this study, we explored “more aggressive” methods for diagnostic purpose by selecting
13
14 relevant wavenumbers. Wavelength selection based on varying window sizes have been reported
15
16 in regression analysis, which depends on the correction of spectrum and the response variables,
17
18 e.g. concentrations of chemicals. We proposed a method to implement variably-sized windows
19
20 as features to prevent over-fitting for discrimination analysis, which depends on the correlation
21
22 between wavenumbers. To our knowledge, this is the first report of wavelength selection based
23
24 on varying window sizes for discrimination analysis. We proved that this method is effective for
25
26 improving skin cancer diagnosis using Raman spectroscopy. It can also be readily used for other
27
28 spectroscopic techniques or diagnosis of other types of cancers.
29
30
31
32
33
34
35

36
37 In order to validate wavenumber selection in Raman spectroscopy for skin cancer diagnosis, two
38
39 types of validation were analyzed. The first analysis was based on independent training and
40
41 testing cohorts. The second analysis was for the combined cohort based on leave-one-out cross-
42
43 validation. Two independent algorithms were applied for wavenumber selection including
44
45 stepwise regression and LASSO, and two independent algorithms were applied for
46
47 discrimination including PC-GDA and PLS. It was found that with wavenumber selection using
48
49 either stepwise regression or LASSO could improve the diagnostic performance including area
50
51 under the ROC curve and specificities at preselected sensitivity levels for all the analyses (**figure**
52
53 **2 – 4 and table 2 – 3**). For cancer diagnosis purpose, high sensitivity is required as missing any
54
55
56
57
58
59
60

1
2
3 cancer cases is fatal to the patients. However, the specificity is relative low at high sensitivity
4
5 levels with our previous algorithms^{1,2}, which causes unnecessary false positive biopsies.
6
7
8 Therefore, one of our goals was to improve the specificity at high sensitivity levels. Our results
9
10 demonstrated that the improvement was significant and clinically relevant with wavenumber
11
12 selection based analysis. For example, the area under the ROC curve was improved from 0.879
13
14 to 0.904-0.905 for training, 0.861 to 0.891-0.911 for testing, and 0.891 to 0.899-0.906 for the
15
16 combined cohort; the specificity at sensitivity level of 90% was improved from 64% to 72-75%
17
18 for training, from 54% to 64-74% for testing, and from 65% to 71% for the combined cohort
19
20 based on PC-GDA analysis (table 3), which represents an improvement of over 10%. Because
21
22 there were fewer cases in the testing cohort (cohort 2), the confident interval for the testing
23
24 cohort was slightly larger than the training cohort (cohort 1) (**figure 2c and 2f, table 2 – 3**). For
25
26 the combined cohort, the confident interval was much smaller than the training cohort and the
27
28 testing cohort (**figure 3, table 2 – 3**).

29
30
31
32
33
34
35
36 In order to assess the contribution of spectral wavebands to the final discrimination, we carried
37
38 PLS analysis on the training cohort and the testing cohort using 100 cm⁻¹ wide wavebands from
39
40 500 to 1800 cm⁻¹ (**figure 1b**). It can be seen that the results of testing cohort are sometimes
41
42 higher or lower than the training cohorts. One possible reason for these variations is that the
43
44 distribution of cases is different for the training cohort and the testing cohort. Another possible
45
46 reason is that the training cohort may be slightly over-fit or under-fit which leads the results of
47
48 the testing cohort lower or higher accordingly. But the results of the training cohort and the
49
50 testing cohort are relatively consistent, and the differences are statistically insignificant ($p >$
51
52 0.05). Figure 1b shows that all the wavebands contributed to the final diagnosis to some extent,
53
54
55
56
57
58
59
60

1
2
3 particularly the waveband $1200\text{-}1400\text{ cm}^{-1}$ provides more diagnostic capability than other
4
5 wavebands, but none of them alone could provide diagnostic performance as high as the full
6
7 wavelength band (**figure 1b** and **Table 2**). Therefore, it was no surprise that the Raman bands
8
9 that were selected using a more rational and rigorous approach was distributed throughout the
10
11 Raman range as measured (**figure 1a**).
12
13
14

15
16
17 Stepwise regression was a classical method, which was widely used in a number of applications.
18
19 However, it had its own limitations⁵⁰. Because it was a localized method, the wavelengths
20
21 selected by stepwise regression were not necessarily globally important. One could not overly
22
23 emphasize the wavelengths found by stepwise regression or LASSO. But if the final purpose was
24
25 discrimination, we found that both stepwise regression and LASSO could be used for
26
27 wavenumber selection to improve diagnosis.
28
29
30
31

32
33
34 Wavelength selection could be implemented based on single wavelengths or by wavelength
35
36 windows. Although wavenumber selection based analyses were reported, there is no report about
37
38 the effect of window sizes. We found that wavelength selection by single wavelength (equivalent
39
40 to window size of 1 pixel or correlation coefficient of 1) tended to over-fit the training cohort,
41
42 and thus the diagnostic performance for the testing cohort was always inferior (See
43
44 supplementary figure S1†). Therefore, one should be very careful when using single wavelength
45
46 selection based analysis. Using windows of wavelength, one could partially overcome the over-
47
48 fitting of the training cohort. In the lung cancer diagnosis by Raman spectroscopy in the high
49
50 wavenumber region, we found that a fixed window size of 5 pixels provided the best diagnostic
51
52 performance, which is equivalent to the spectrometer resolution⁴⁰. However, because fixed-size
53
54
55
56
57
58
59
60

1
2
3 windows might split the correlated wavelength into different windows, the improvement with
4
5 fixed window sizes was not consistent (See supplementary figure S1†). Here we applied a simple
6
7 method to implement varying window sizes based on formal correlation analysis. With the
8
9 proposed varying size of windows as extracted using either of two wavenumber selection
10
11 methods, the performance of the training cohort, the testing cohort and the combined cohort were
12
13 all improved. The computational load of calculating correlation between wavenumber is minimal,
14
15 and thus it is worth of the extra computation to implement variably-size windows than fixed-size
16
17 windows in wavenumber-based analysis.
18
19
20
21
22
23
24

25 26 **5. Conclusions**

27
28
29 In summary, we explored wavenumber selection based analysis with variably-sized windows of
30
31 wavenumbers in Raman spectroscopy for *in vivo* skin cancer diagnosis. We found that
32
33 wavenumber selection based analysis could substantially improve the diagnostic specificity.
34
35 With LASSO-based wavenumber selection and PC-GDA analysis, the area under the receiver
36
37 operating characteristic curve (ROC) was improved from 0.861 – 0.891 to 0.891 – 0.911 and the
38
39 diagnostic specificity was improved from 0.17 – 0.65 to 0.20 – 0.75 for fixed sensitivities
40
41 between 0.99 – 0.90. Analysis based on stepwise regression and PLS showed similar
42
43 improvement in diagnostic performance.
44
45
46
47
48
49

50 51 **Acknowledgment**

52
53 The project is financially supported by the Canadian Cancer Society (CCS), the Canadian
54
55 Institutes of Health Research (CIHR), the Canadian Dermatology Foundation (CDF), the VGH &
56
57

1
2
3 UBC Hospital Foundation, the BC Hydro Employees Community Service Fund and in part by
4
5 the Verisante Technology Inc. Dr. David McLean, Wei Zhang, Mohammed AlJasser and
6
7
8 Soodabeh Zandi are greatly acknowledged for their contributions in consultation, patient
9
10 recruitment and clinical measurement.
11

12 13 14 15 **Conflict of Interest**

16
17 J. Zhao, H. Lui, H. Zeng, and the BC Cancer Agency hold patent rights in Raman spectroscopy
18
19 that are licensed to Verisante Technology Inc.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Case summary of the lesions in the analysis. Cohort one was used for training and cohort two was used for testing reported by Lui et al ¹ and Zhao et al ². The combined cohort was a combination of cohort one and cohort two for leave-one-out cross-validation analysis. SCC: squamous cell carcinoma; BCC: basal cell carcinoma.

Final Lesion Diagnosis		Cohort One	Cohort Two	Combined Cohort
Melanoma	Lentigo maligna	20	0	20
	Lentigo maligna melanoma	8	2	10
	Superficial spreading melanoma	14	7	21
	Melanoma other type	2	0	2
BCC	BCC superficial	28	6	34
	BCC nodular	73	24	97
	BCC other type	8	5	13
SCC	SCC in situ	18	5	23
	SCC invasive	28	13	41
	SCC other type	1	0	1
Actinic keratosis		32	15	47
Atypical nevus		57	16	73
Blue nevus		13	0	13
Compound nevus		30	2	32
Intradermal nevus		38	7	45
Junctional nevus		34	1	35
Seborrheic keratosis		114	24	138
Total		518	127	645

Table 2. Diagnostic performance of Raman spectroscopy according to model, including the area under the ROC curve and the 95% confidence intervals in parenthesis. The models consisted of wavelength selection (full band, LASSO or stepwise regression) and diagnosis (PC-GDA or PLS). PC-GDA: principal component and general discrimination analysis, PLS: partial least squares, LASSO: least and shrinkage selection operator, LOOCV: leave-one-out cross-validation, CV: cross-validation.

Model	Training cohort (n = 518, LOOCV)	Testing cohort (n = 127, CV)	Combined cohort (n = 645, LOOCV)
Full band, PC-GDA	0.879 (0.829 – 0.929)	0.861 (0.796 – 0.927)	0.891(0.867 – 0.916)
LASSO, PC-GDA	0.905 (0.879 – 0.931)	0.891 (0.835 – 0.948)	0.906(0.883 – 0.929)
Stepwise, PC-GDA	0.904 (0.878 – 0.930)	0.911 (0.863 – 0.959)	0.899 (0.876 – 0.922)
Full band, PLS	0.896 (0.846 – 0.946)	0.889 (0.834 – 0.944)	0.894(0.870 – 0.918)
LASSO, PLS	0.909 (0.884 – 0.935)	0.895 (0.839 – 0.950)	0.907(0.884 – 0.929)
Stepwise, PLS	0.912 (0.887 – 0.936)	0.904 (0.853 – 0.954)	0.907 (0.885 – 0.929)

Table 3. Summary of Raman spectroscopy diagnostic specificity and the 95% confidence intervals in parenthesis derived from ROCs according to various levels of sensitivity. The models consisted of wavenumber selection (full band, LASSO or stepwise regression) and diagnosis (PC-GDA or PLS). PC-GDA: principal component and general discrimination analysis, PLS: partial least squares, LASSO: least and shrinkage selection operator, LOOCV: leave-one-out cross-validation, CV: cross-validation.

Model	Sensitivity	Specificity		
		Training set (n = 518, LOO-CV)	Test set (n = 127 CV)	Combined set (n = 645, LOO-CV)
Full band, PC-GDA	0.99 (0.96 – 1.00)	0.17 (0.13 – 0.21)	0.24 (0.13 – 0.38)	0.24 (0.20 – 0.29)
	0.95 (0.91 – 0.97)	0.41 (0.35 – 0.48)	0.48 (0.34 – 0.63)	0.54 (0.48 – 0.59)
	0.90 (0.85 – 0.93)	0.64 (0.58 – 0.70)	0.54 (0.39 – 0.68)	0.65 (0.60 – 0.70)
LASSO PC-GDA	0.99 (0.96 – 1.00)	0.27 (0.22 – 0.33)	0.20 (0.10 – 0.34)	0.26 (0.21 – 0.31)
	0.95 (0.91 – 0.97)	0.55 (0.49 – 0.61)	0.48 (0.34 – 0.63)	0.55 (0.50 – 0.61)
	0.90 (0.85 – 0.93)	0.72 (0.67 – 0.77)	0.74 (0.60 – 0.85)	0.71 (0.65 – 0.75)
Stepwise, PC-GDA	0.99 (0.96 – 1.00)	0.33 (0.28 – 0.39)	0.34 (0.21 – 0.49)	0.35 (0.26 – 0.40)
	0.95 (0.91 – 0.97)	0.56 (0.50 – 0.61)	0.54 (0.39 – 0.68)	0.54 (0.48 – 0.59)
	0.90 (0.85 – 0.93)	0.75 (0.70 – 0.80)	0.64 (0.49 – 0.77)	0.71 (0.66 – 0.76)
Full band, PLS	0.99 (0.96 – 1.00)	0.24 (0.19 – 0.29)	0.30 (0.18 – 0.45)	0.24 (0.19 – 0.28)
	0.95 (0.91 – 0.97)	0.52 (0.48 – 0.58)	0.46 (0.32 – 0.61)	0.54 (0.48 – 0.59)
	0.90 (0.85 – 0.93)	0.66 (0.61 – 0.71)	0.62 (0.47 – 0.75)	0.67 (0.62 – 0.72)
LASSO, PLS	0.99 (0.96 – 1.00)	0.25 (0.20 – 0.31)	0.30 (0.18 – 0.45)	0.27 (0.22 – 0.32)
	0.95 (0.91 – 0.97)	0.58 (0.52 – 0.64)	0.44 (0.30 – 0.59)	0.58 (0.53 – 0.63)
	0.90 (0.85 – 0.93)	0.76 (0.71 – 0.81)	0.72 (0.58 – 0.84)	0.73 (0.68 – 0.78)
Stepwise, PLS	0.99 (0.96 – 1.00)	0.38 (0.33 – 0.44)	0.30 (0.18 – 0.45)	0.33 (0.28 – 0.38)
	0.95 (0.91 – 0.97)	0.62 (0.56 – 0.67)	0.56 (0.41 – 0.70)	0.59 (0.54 – 0.64)
	0.90 (0.85 – 0.93)	0.74 (0.69 – 0.79)	0.72 (0.58 – 0.84)	0.73 (0.68 – 0.78)

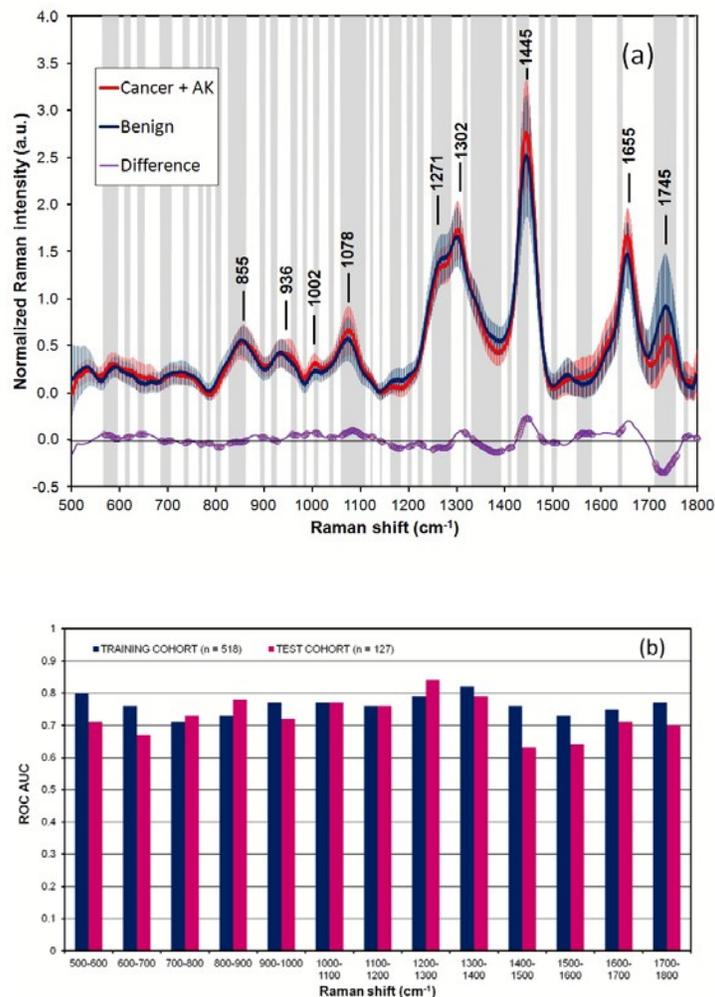


Figure 1 (a) Averaged Raman spectra and standard deviation of skin cancers (including precancers) and benign skin lesions of the combined cohort ($n = 309$ cancer vs $n = 336$ benign). Each spectrum is normalized to their respective area under the curve between 500 and 1800 cm^{-1} before averaging. The difference spectrum between skin cancers and benign skin lesions is also shown (shifted by 0.5 unit for illustration purpose). The circles on the difference spectrum and the shaded area represent the selected wavebands from LASSO based on LOO-CS of the training cohort only. (b) The diagnostic performance of each Raman band spanning 100 cm^{-1} wide from 500 cm^{-1} to 1800 cm^{-1} for the training cohort and the testing cohort based on PLS. Note that all the Raman bands contribute to the discrimination between skin cancers and benign skin lesions, but no single Raman band provides diagnosis performance as high as the full Raman band.

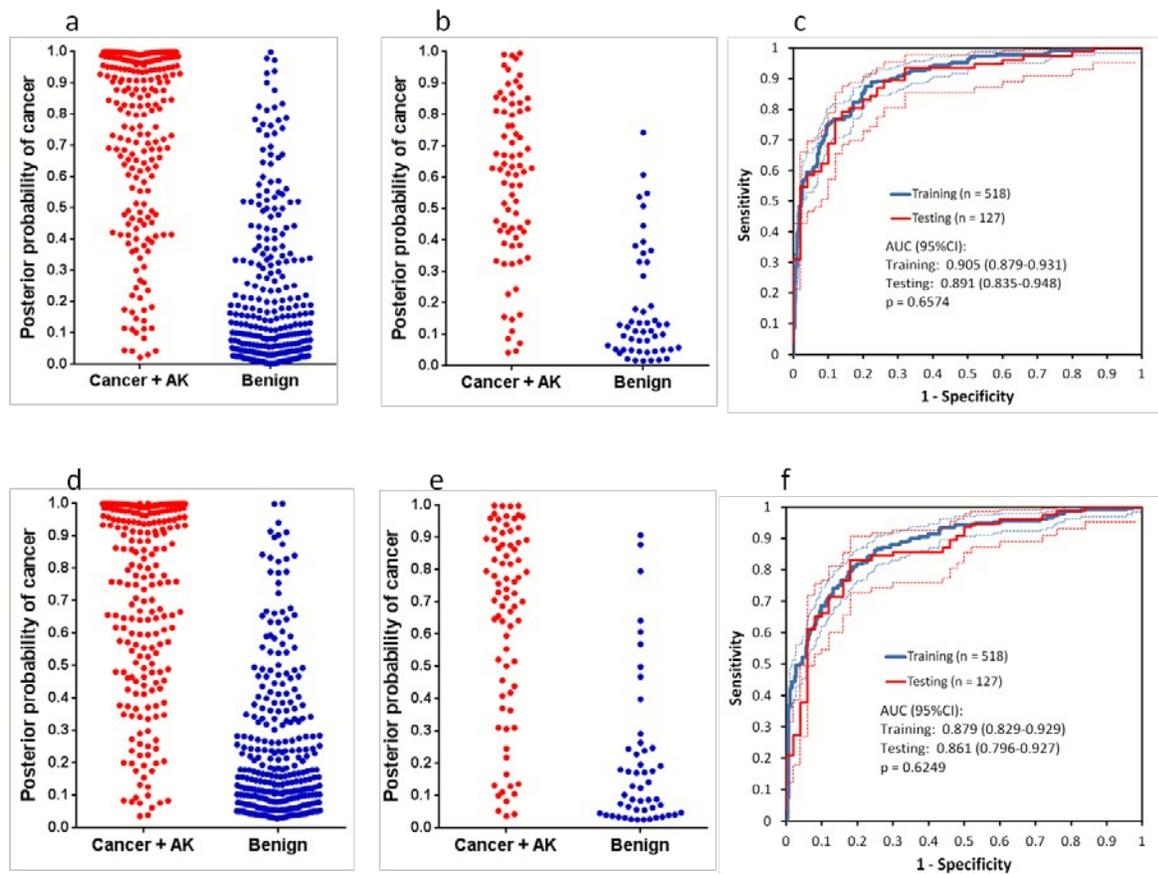


Figure 2. Lesion classification by Raman spectroscopy for the training cohort ($n = 518$) and testing cohort ($n = 127$) by PC-GDA analysis with wavenumber selection (a-c) and without wavenumber selection (d-f). (a,b) and (d,e) are the posterior probabilities for discriminating skin cancers and precancers from benign skin disorders for the training cohort (a) and the testing cohort (b) based on wavelength selection by LASSO, and for the training cohort (d) and the testing cohort (e) based on the full wavelength band. (c, f) are the corresponding ROC curves (solid lines) and 95% CIs (dashed lines) derived from the respective posterior probabilities in (a, b) and (d, e) respectively. Note that there is no statistical difference between the training cohort and the testing cohort for both the full wavelength band ($p = 0.6249$) and the selected Raman bands ($p = 0.6574$).

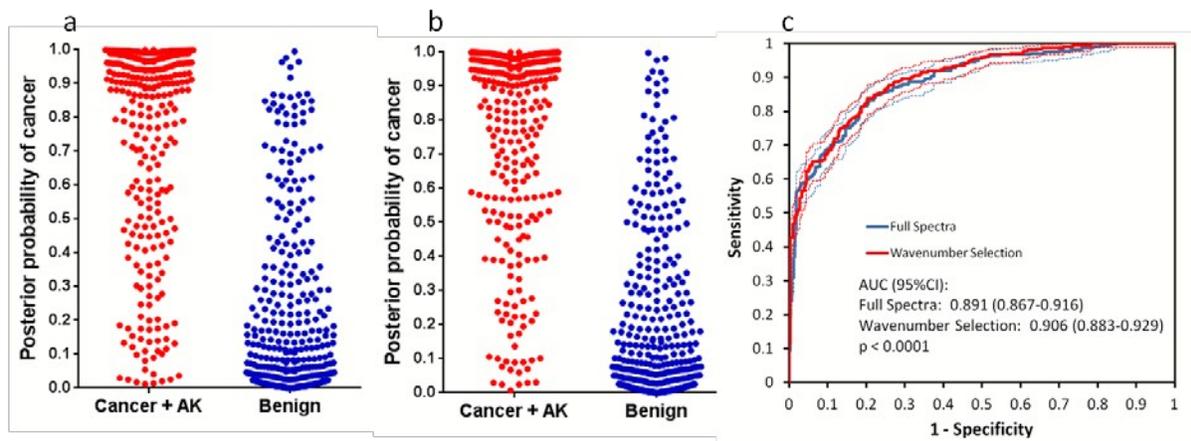


Figure 3. Lesion classification by Raman spectroscopy for the combined cohort ($n = 645$) by PC-GDA without wavenumber selection (a) and with wavenumber selection (b). Posterior probabilities for discriminating skin cancers and precancers from benign skin disorders for the full wavelength band (a) and for the wavelength selection by LASSO (b). (c) The corresponding ROC curves (solid lines) and 95% CIs (dashed lines) are derived from the respective posterior probabilities in (a, b). Note that the AUC is improved from 0.891 to 0.906 after wavenumber selection and the improvement is statistically significant $p < 0.0001$.

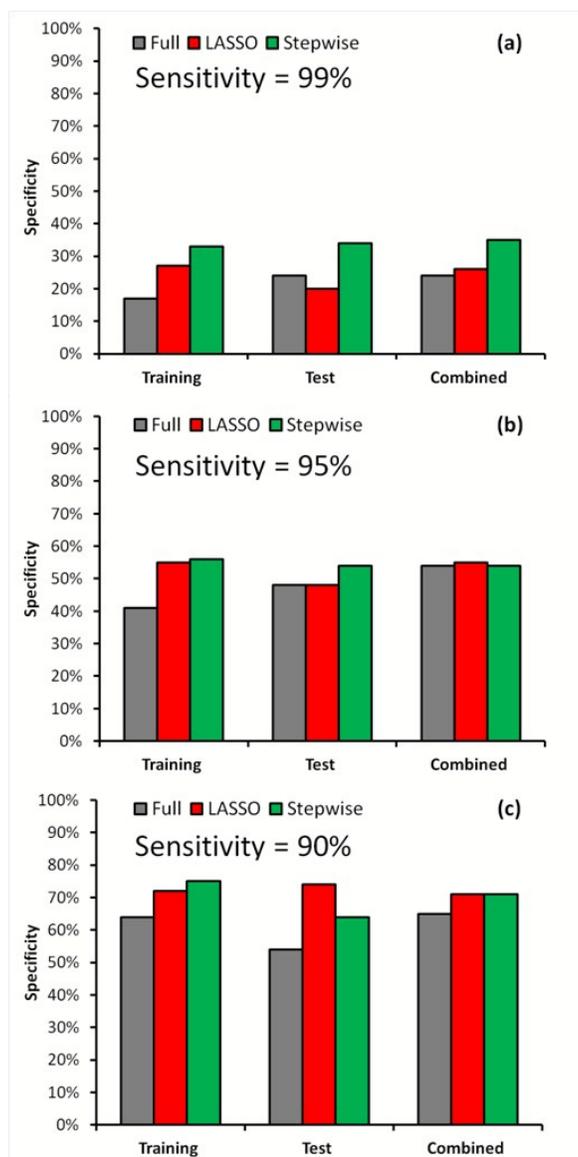


Figure 4. Diagnostic specificity based on PC-GDA at various sensitivity levels for the training cohort, the test cohort and the combined cohort with and without wavenumber selection based analysis. (a) sensitivity = 99%, (b) sensitivity = 95%, (c) sensitivity = 90%.

References

1. H. Lui, J. Zhao, D. McLean and H. Zeng, *Cancer Res*, 2012, **72**, 2491-2500.
2. J. Zhao, H. Lui, S. Kalia and H. Zeng, *Anal Bioanal Chem*, 2015, DOI: 10.1007/s00216-015-8914-9.
3. J. Zhao, H. Lui, D. I. McLean and H. Zeng, *Skin Res Technol*, 2008, **14**, 484-492.
4. Z. Huang, H. Zeng, I. Hamzavi, D. I. McLean and H. Lui, *Opt Lett*, 2001, **26**, 1782-1784.
5. C. A. Lieber, S. K. Majumder, D. Billheimer, D. L. Ellis and A. Mahadevan-Jansen, *J. Biomed. Opt.*, 2008, **13**, 024013.
6. C. A. Lieber, S. K. Majumder, D. L. Ellis, D. Billheimer and A. Mahadevan-Jansen, *Lasers Surg Med*, 2008, **40**, 461-467.
7. A. Nijssen, K. Maquelin, L. F. Santos, P. J. Caspers, T. C. Bakker Schut, J. C. den Hollander, M. H. A. Neumann and G. J. Puppels, *J Biomed Opt*, 2007, **12**, 034004.
8. Z. Huang, A. McWilliams, H. Lui, D. I. McLean, S. Lam and H. Zeng, *Int J Cancer*, 2003, **107**, 1047-1052.
9. M. Short, S. Lam, A. McWilliams, J. Zhao, H. Lui and H. Zeng, *Opt. Lett.*, 2008, **33**, 711-713.
10. M. A. Short, S. Lam, A. M. McWilliams, D. N. Ionescu and H. Zeng, *J Thorac Oncol*, 2011, **6**, 1206-1214.
11. M. A. Short, I. T. Tai, D. Owen and H. Zeng, *Opt Express*, 2013, **21**, 5025-5034.
12. M. S. Bergholt, W. Zheng, K. Lin, J. Wang, H. Xu, J.-L. Ren, K. Y. Ho, M. Teh, K. G. Yeoh and Z. Huang, *Anal Chem*, 2015, **87**, 960-966.
13. E. Widjaja, W. Zheng and Z. Huang, *Int. J. Oncology*, 2008, **32**, 653-662.
14. M. S. Bergholt, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So, A. Shabbir and Z. Huang, *J Biophotonics*, 2013, **6**, 49-59.
15. M. S. Bergholt, W. Zheng, K. Lin, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So and Z. Huang, *Biosens Bioelectron*, 2011, **26**, 4104-4110.
16. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, U. Utzinger and R. Richards-Kortum, *Photochem. Photobiol.*, 1998, **68**, 427-431.
17. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, A. Malpica, S. Thomsen, U. Utzinger and R. Richards-Kortum, *Photochem Photobiol*, 1998, **68**, 123-132.
18. U. Utzinger, D. L. Heintzelman, A. Mahadevan-Jansen, A. Malpica, M. Follen and R. Richards-Kortum, *Appl. Spectrosc.*, 2001, **55**, 955-959.
19. J. Mo, W. Zheng, J. J. H. Low, J. Ng, A. Ilancheran and Z. Huang, *Anal Chem*, 2009, **81**, 8908-8915.
20. S. Duraipandian, W. Zheng, J. Ng, J. J. H. Low, A. Ilancheran and Z. Huang, *Analyst*, 2011, **136**, 4328-4336.
21. K. Guze, H. C. Pawluk, M. Short, H. Zeng, J. Lorch, C. Norris and S. Sonis, *Head & neck*, 2014, **37**, 511-517.
22. K. Guze, M. Short, S. Sonis, N. Karimbux, J. Chan and H. Zeng, *J Biomed Opt*, 2009, **14**, 014016.
23. M. V. P. Chowdary, K. K. Kumar, J. Kurien, S. Mathew and C. M. Krishna, *Biopolymers*, 2006, **83**, 556-569.
24. M. V. P. Chowdary, K. K. Kumar, S. Mathew, L. Rao, C. M. Krishna and J. Kurien, *Biopolymers*, 2009, **91**, 539-546.
25. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari and M. S. Feld, *Proc Natl Acad Sci U S A*, 2005, **102**, 12371-12376.
26. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R. R. Dasari, J. P. Crowe and M. S. Feld, *Cancer Res*, 2006, **66**, 3317-3322.

- 1
- 2
- 3
- 4 27. A. S. Haka, Z. Volynskaya, J. A. Gardecki, J. Nazemi, R. Shenk, N. Wang, R. R. Dasari, M. Fitzmaurice and M. S. Feld, *J Biomed Opt*, 2009, **14**.
- 5
- 6 28. M. D. Keller, E. Vargis, N. de Matos Granja, R. H. Wilson, M.-A. Mycek, M. C. Kelley and A. Mahadevan-Jansen, *J Biomed Opt*, 2011, **16**, 077006.
- 7
- 8 29. N. Stone and P. Matousek, *Cancer Res*, 2008, **68**, 4424-4430.
- 9 30. A. Hoskuldsson, *Chemometr Intell Lab*, 2001, **55**, 23-38.
- 10 31. L. Norgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck and S. B. Engelsen, *Appl Spectrosc*, 2000, **54**, 413-419.
- 11 32. E. Anderssen, K. Dyrstad, F. Westad and H. Martens, *Chemometr Intell Lab*, 2006, **84**, 69-74.
- 12 33. C. M. Andersen and R. Bro, *Journal of Chemometrics*, 2010, **24**, 728-737.
- 13 34. Zoran Bursac, C Heath Gauss, D. K. Williams and D. W. Hosmer, *Source Code for Biology and Medicine*, 2008, **2008**, 17.
- 14
- 15 35. A. Talukder and D. Casasent, *Opt Eng*, 1998, **37**, 904-913.
- 16 36. S. K. Majumder, M. D. Keller, F. I. Boulos, M. C. Kelley and A. Mahadevan-Jansen, *J Biomed Opt*, 2008, **13**.
- 17 37. S. K. Majumder, N. Ghosh, S. Kataria and P. K. Gupta, *Lasers Surg Med*, 2003, **33**, 48-56.
- 18 38. S. K. Majumder, S. Gebhart, M. D. Johnson, R. Thompson, W.-C. Lin and A. Mahadevan-Jansen, *Appl Spectrosc*, 2007, **61**, 548-557.
- 19 39. S. Li, L. Li, Q. Zeng, Y. Zhang, Z. Guo, Z. Liu, M. Jin, C. Su, L. Lin and J. Xu, *Sci Rep*, 2015, **5**.
- 20 40. H. C. McGregor, M. A. Short, A. McWilliams, T. Shaipanich, D. N. Ionescu, J. Zhao, W. Wang, G. Chen, S. Lam and H. Zeng, *J. Biophotonics*, 2015, (In Press, DOI: 10.1002/jbio.201500204).
- 21 41. J. Zhao, H. Lui, D. I. McLean and H. Zeng, *Appl Spectrosc*, 2007, **61**, 1225-1232.
- 22 42. R. Leardi and A. L. Gonzalez, *Chemometr Intell Lab*, 1998, **41**, 195-207.
- 23 43. R. Leardi, *Journal of Chemometrics*, 2000, **14**, 643-655.
- 24 44. M. Forina, S. Lanteri, M. C. C. Oliveros and C. P. Millan, *Anal Bioanal Chem*, 2004, **380**, 397-418.
- 25 45. I. Guyon and A. Elisseeff, *The Journal of Machine Learning Research*, 2003, **3**, 1157-1182.
- 26 46. R. Tibshirani, *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, 267-288.
- 27 47. T. Hill and P. Lewicki, *Statistics: Methods and Applications*, Statsoft, Tulsa, OK, USA, 2005.
- 28 48. J. A. Hanley and B. J. McNeil, *Radiology*, 1982, **143**, 29-36.
- 29 49. J. A. Hanley and B. J. McNeil, *Radiology*, 1983, **148**, 839-843.
- 30 50. D. L. Streiner, *Can. J. Psychiatry*, 1994, **39**, 191.
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60