

FTIR spectroscopy of lung cells on glass coverslips. A further step towards spectral pathology?

Background

Cancer incidence rates have been continually rising since the early 1990s [1]. This is putting further pressure on pathology departments which can delay diagnosis and worsen patient outcomes. This has only been aggravated by backlogs from COVID-19. An automated system using Fourier transform infrared (FTIR) spectroscopy could help to improve diagnosis times while providing an objective diagnosis. A major hurdle to the translation of FTIR spectroscopy to a clinical setting is the cost of substrates. Glass is not regularly used as a substrate because the fingerprint region of the spectra is obscured. Using thinner glass allows information on lipids and proteins to be seen [2, 3, 4]. We have shown that enough information is available in the spectra to classify lung cancer cells from non-malignant lung cells using a glass substrate.

Methods

Spectra acquisition

Cell lines

Lung cancer: A549, CALU-1
Non-malignant lung: NL20

Sample preparation

Cells applied to glass coverslip (0.15-0.17mm) by cytopsin and fixed with 4% PFA. 3 independent experiments in triplicate.

Measurements

180 spectra of each cell line collected from individual cells using synchrotron transmission FTIR spectroscopy.

Classification

Train/Test

Spectra were randomly sampled. Split 2/3 for training and 1/3 for testing. Repeated 100 times.

KNN

Neighbours = 4
Distance = Euclidean

Random forest

number of trees = 100

Pre-processing

Cropping

Spectra cropped to two regions:
3100-2700 cm^{-1} and 1800-1350 cm^{-1}

S-G filter

Window = 21
Polynomial = 2
Derivative = 2

Normalisation

SNV

PCA

Retained 95% of explained variance.

Conclusions

- A glass substrate combined with machine learning is viable for classifying lung cancer cells from non-malignant lung cells with FTIR spectroscopy.
- The lipid bands (3100-2700 cm^{-1}) provided the best classification for both A549 and CALU-1 cell lines from NL20 cell line.
- The glass substrates are far more affordable than the CaF_2 and BaF_2 slides currently used.

References

- Cancer-statistics/incidence#heading=Zero. Published 2018. Accessed August 12, 2020.
- Rutter A V., Crees J., Wright H., van Pittius DG, Yousef I, Sulé-Suso J. Fourier transform infrared spectra of cells on glass coverslips. A further step in spectral pathology. *Analyst*. 2018;143(23):5711-5717. doi:10.1039/C8AN01634H
- Rutter A V, Crees J, Wright H, et al. Identification of a Glass Substrate to Study Cells Using Fourier Transform Infrared Spectroscopy: Are We Closer to Spectral Pathology? *Appl Spectrosc*. 2019. doi:10.1177/0003702819875828
- Dowling L, Roach P, Rutter A V, et al. Optimization of Sample Preparation Using Glass Slides for Spectral Pathology. *Appl Spectrosc*. doi:10.1177/0003702820945748

Glass substrates combined with machine learning can be used to classify lung cancer cells from non-malignant lung cells with FTIR spectroscopy.

Results

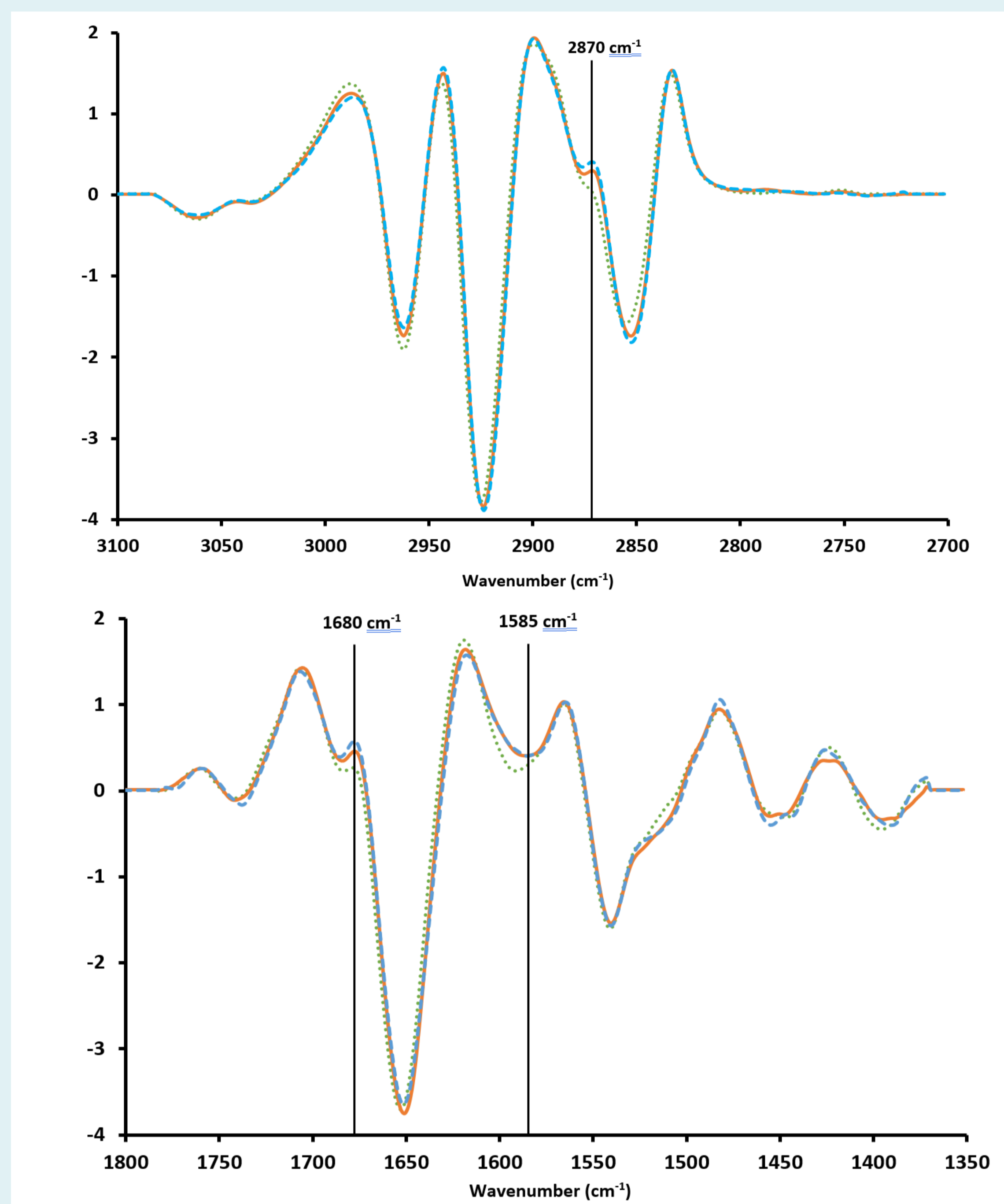


Figure 1. Average 2nd derivative spectra of NL20 (dotted), CALU-1 (dashed), A549 (solid) in the regions 3100-2700 cm^{-1} (top) and 1800-1350 cm^{-1} (bottom).

Cells	Region of spectra	Classifier	AUC	CA	Sensitivity (%)	Specificity (%)
A549 vs NL20	Region between 1800 cm^{-1} to 1350 cm^{-1} (amide I & II)	RF	0.919	0.836	93.2	72.3
		KNN	0.897	0.829	85.1	80.1
A549 vs NL20	Region between 3100 cm^{-1} to 2700 cm^{-1} (lipids)	RF	0.967	0.916	91.9	91.0
		KNN	0.935	0.891	93.0	84.4
CALU-1 vs NL20	Region between 1800 cm^{-1} to 1350 cm^{-1} (amide I & II)	RF	0.944	0.865	92.5	80.2
		KNN	0.912	0.830	78.0	88.2
CALU-1 vs NL20	Region between 3100 cm^{-1} to 2700 cm^{-1} (lipids)	RF	0.992	0.953	95.1	94.2
		KNN	0.971	0.940	94.6	84.4

Table 1. Random forest (RF) and K-nearest neighbors (KNN) classification models for A549 vs NL20 and CALU-1 vs NL20 using the spectral regions 3100-2700 cm^{-1} and 1800-1350 cm^{-1} .

