

Assisted Real-Time Label-Free SERS by Machine Learning Malignant Pleural Effusion Diagnoses Associated with Lung Cancer

Pyng Olivo*

Department of Respiratory and Critical Care Medicine, National University Hospital, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 11759, Singapore

Introduction

One of the numerous diseases with high death rates is cancer. As the world's population ages and adopts unhealthy lifestyles, the prevalence of cancer is rising, placing a tremendous financial burden on healthcare systems, particularly in poorer nations. As of 2012, lung cancer in particular was the most frequently diagnosed cancer, accounting for 13% of all cancer occurrences and 20% of all cancer deaths. With one of the lowest 5-year survival rates of 10–15%, it is still a fatal illness. Clinical stage IV lung cancer patients had a 5-year overall survival rate of just 2% between 1975 and 2009, compared to a 5-year overall survival rate of 50% for patients with the lowest stage IA. The usual variety of exams, including physical examinations, imaging tests, and biopsies—all of which can be invasive and time-consuming—are part of the diagnostic process. For a positive patient prognosis, precise, quick, and non-invasive diagnostic techniques are essential for the early diagnosis of lung cancer. As it allows for the spectral study of biofluids and tissues in the body, Raman spectroscopy is developing as a promising practical option for the diagnosis of illnesses, including cancer.

This method has been used in several studies to try and pinpoint biomolecular and chemical alterations in biofluids, such as serum, plasma, saliva, cervical fluid, and urine, that are connected to cancer. The method of chemical information, or chemometrics, is useful and required to extract meaningful information and conduct systematic study on the gathered spectral data. In order to produce SERS spectra, we have built a dependable and high level of homogeneity throughout the SERS substrate. For each sample of pleural fluid, we have utilised mapping to obtain numerous SERS spectra. The SCSNP SERS substrate and chemometric pattern recognition techniques, such as principal component analysis (PCA), linear correlation analysis (LDA), partial least squares-discriminant analysis (PLS-DA), and uncorrelated linear discriminant analysis, were used to distinguish between controls and cancer patients (ULDA). The potential of the SERS platform combined with the application of a binary classifier algorithm for the quick diagnosis of lung cancer is confirmed by our proof-of-concept study with a limited sample size with high sensitivity and specificity using pleural fluid [1].

Description

To separate the fluid from the cells, pleural fluid samples were placed in 1.5 mL Eppendorf™ tubes and centrifuged at 10,000 rpm for 10 min at 4°C. A portion of the supernatant was taken off, aliquoted, and kept at

20°C. For measurement, only specimens that showed no traces of blood contamination were employed. To improve the wettability of aqueous samples, these substrates were cleaned with ethanol before being used, followed by Millipore water. Five litres of frozen pleural fluid were defrosted and applied to the substrate. Using a 785 nm laser and a Renishaw InVia Raman upright microscope (Renishaw InVia, Gloucestershire, UK), SERS measurements of the pleural fluid were made. The Raman system was paired with a Leica microscope, and the laser light was transmitted through an objective lens (50X, 0.75 N.A), which was utilised to stimulate the sample and gather the Raman signal that was scattered. A notch filter was used to block the predominant Rayleigh scattering, and the beam point on the sample was around 1 m [2].

Per pleural fluid sample, at least 500 spectra were obtained through mapping over three separate sites. Each spectrum's integration duration was 10 seconds, and the resulting mapping data were gathered in the 700–1800 cm^{-1} wavenumber band. The concept for measuring the pleural fluid using label-free SERS on a SCSNP substrate. Our earlier work offers a description of how the SCSNP was constructed. To put it briefly, the Si was etched using a blanket inductive coupled plasma (ICP) etching process. E-beam evaporation is used to deposit pure silver (Ag) after creating a silicon-based nanograin structure at random on a wafer. In this pilot investigation, we used benchtop confocal Raman spectroscopy to get the SERS spectra from pleural fluid samples of people with various illnesses (lung, breast, ovarian cancer, etc.). The pleural fluid sample's Raman fingerprint, which was obtained between 735 cm^{-1} and 1700 cm^{-1} , was also employed in the chemometrics study. Using the associated pleural fluid SERS sample as a mapping reference, several spectra were obtained for each patient [3].

When the sample was irradiated with the laser light source in the confocal Raman spectroscopy system, autofluorescence from the sample had an impact on the mapping data in terms of numerous SERS spectra. The sample's autofluorescence was overlapping. SERS signal strength and the corresponding molecule concentration were not linearly correlated with the SERS spectral signal from the sample under investigation. The SERS spectra have to be preprocessed in order to eliminate any autofluorescence baseline effects. The asymmetric least squares (AsLS) baseline correction approach was used to build a spectral baseline correction in order to mathematically eliminate the effect of autofluorescence. In order to reduce cosmic ray spike contributions and improve the signal-to-noise ratio (SNR), the baseline-corrected SERS spectra were further processed using the Savitzky-Golay smoothing technique. For each SERS spectrum acquired for each pleural fluid sample, the preprocess step was repeated. The subject's SERS spectrum after baseline correction with increased SNR was divided further into the control and cancer groups, and then subjected to a binary classification analysis using the partial least squares-discriminant analysis (PLS-DA) method [4].

For dimensionality reduction and additional categorization, we constructed a supervised PLS-DA based on partial least squares regression (PLSR). In comparison to the PCA model, the PLS regression-based PLS-DA binary classification method further rotated the components (latent variables, or LVs), resulting in superior classification outcomes. The PLS-DA classification method's fundamental theory has been discussed elsewhere. A stratified K (10)-fold cross-validation method was employed to validate the precision of this PLS-DA binary classification model, which helped to further reduce classification mistakes brought on by class imbalance. The PLS-DA classification model can

*Address for Correspondence: Pyng Olivo, Department of Respiratory and Critical Care Medicine, National University Hospital, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore; E-mail: Pyngolivo@.edu.sg

Copyright: © 2022 Olivo P. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Date of Submission: 01 September, 2022, Manuscript No. LDT-22-81082; Editor Assigned: 03 September, 2022, PreQC No. P-81082; Reviewed: 17 September, 2022, QC No. Q-81082; Revised: 23 September, 2022, Manuscript No. R-81082; Published: 01 October 2022, DOI: 10.37421/2472-1018.2022.8.164

be used to evaluate the variable importance in projection (VIP) score, which estimates the significance of each variable. The weighted VIP score is the PLS latent variable summation correlation. The PLS-DA classification model can be further optimised by taking into account variables that have a VIP score value larger than one. The VIP score can be further used to differentiate between the two subject groups by choosing certain wavenumbers, as it aids in identifying significant wavenumbers or Raman band regions that are different in the two groups under inquiry. In this investigation, pleural fluid samples from three distinct classes—lung cancer, all other cancers, and control subjects—were examined. With the help of the powerful machine learning algorithms mentioned above, we investigated how the pleural fluid SERS signature could help distinguish between various health conditions [5].

Conclusion

The potential of the SERS platform in conjunction with the application of a binary classifier algorithm for the quick identification of lung cancer using pleural fluid is confirmed by this proof-of-concept study with a modest sample size. In order to determine the disease condition in our current work, we used chemometrics techniques such as principal component analysis followed by linear discrimination analysis (PCA-LDA) and partial least squares-discriminant analysis (PLS-DA), among others. By applying the PLS-DA binary classifier to differentiate between lung tumours and control individuals, we demonstrate a classification accuracy of 85%, a sensitivity of 87%, and specificities of

83% for the identification of lung cancer above control pleural fluid samples. Furthermore, we assessed the discriminative wavenumber. Using the variable significance in projection (VIP) score, bands are in charge of differentiating between the two classes. Our proof-of-concept work shows that it is possible to quickly detect the existence of lung cancer using only the patients' pleural effusions by combining the label-free SERS methodology with cutting-edge chemometrics techniques and the reliable and repeatable SERS substrate (SCSNP).

References

1. Torre, Lindsey A., Freddie Bray, Rebecca L. Siegel and Ahmedin Jemal et al. "Global cancer statistics, 2012." *Ca-Cancer J Clin* 65 (2015): 87-108.
2. McGuire, Shelley. "World cancer report 2014. Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO Press, 2015." *Adv Nutr* 7 (2016): 418-419.
3. Choi, Samjin, Hun-Kuk Park, Gyeong Eun Min and Yeon-Hee Kim. "Biochemical investigations of human papillomavirus-infected cervical fluids." *Microsc Res Tech* 78 (2015): 200-206.
4. Nie, Shuming and Steven R. Emory. "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering." *sci* 275 (1997): 1102-1106.
5. Perumal, Jayakumar, Kien Voon Kong, U. S. Dinis and Malini Olivo et al. "Design and fabrication of random silver films as substrate for SERS based nano-stress sensing of proteins." *RSC Adv* 4 (2014): 12995-13000.

How to cite this article: Olivo, Pyng. "Assisted Real-Time Label-Free SERS by Machine Learning Malignant Pleural Effusion Diagnoses Associated with Lung Cancer." *J Lung Dis Treat* 8 (2022): 164.