

SUPPLEMENTARY MATERIAL

Identification of pulmonary edema in forensic autopsy cases of fatal anaphylactic shock using Fourier transform infrared microspectroscopy

International Journal of Legal Medicine

Hancheng Lin^{1,2&} • Yiwen Luo^{1&} • Lei Wang¹ • Kaifei Deng¹ • Qiran Sun¹ • Ruoxi Fang² • Xin Wei² • Shuai Zha² • Zhenyuan Wang^{2*} • Ping Huang^{1*}

¹ Shanghai Key Laboratory of Forensic Medicine, Shanghai Forensic Service Platform, Institute of Forensic Science, Ministry of Justice, Shanghai, 200063, China

² Department of Forensic Pathology, Xi'an Jiaotong University, Xi'an, 710061, China

[&] These authors contributed equally to this work.

^{*} Corresponding authors

Zhenyuan Wang (E-mail: wzy218@xjtu.edu.cn, Tel: +86 029-82655472)

Ping Huang (E-mail: huangp@ssfjd.cn, Tel: +86 021-52367986)

Table S1. Case profiles

Number	Cause of death	Gender	Age, years	PMI, hours	Allergens
1	Fatal anaphylactic shock	Female	45	<12	Cephalosporin
2	Fatal anaphylactic shock	Male	34	<12	Cephalosporin
3	Fatal anaphylactic shock	Female	73	<12	Cephalosporin
4	Fatal anaphylactic shock	Female	37	<12	Cephalosporin
5	Fatal anaphylactic shock	Male	73	<12	Bone peptide
6	Fatal anaphylactic shock	Male	27	<24	Cephalosporin
7	Fatal anaphylactic shock	Female	52	<12	Lincomycin
8	Fatal anaphylactic shock	Female	25	<6	Cephalosporin
9	Mechanical asphyxia	Female	42	<48	-
10	Mechanical asphyxia	Female	35	<12	-
11	Mechanical asphyxia	Male	16	<24	-
12	Brain injury	Male	27	<24	-
13	Brain injury	Male	30	<24	-
14	Brain injury	Male	48	<24	-
15	Brain injury	Male	19	<24	-
16	Brain injury	Male	40	<12	-
17	Acute cardiac death	Female	55	<12	-
18	Acute cardiac death	Male	65	<12	-
19	Acute cardiac death	Male	53	<12	-
20	Acute cardiac death	Female	81	<12	-

Table S2. The main postmortem investigation results pertaining to the cases

	Medical history and circumstance of death	Necroscopic investigation		Laboratory investigation
		Macroscopic investigation	Histological investigation	
Fatal anaphylactic shock (n = 8)	Previous allergic reactions to certain drugs (8/8) Exposure to an allergen prior to death (8/8) Clinical presentation of an allergic reaction (8/8)	Pharyngeal/laryngeal edema (8/8) Mucous plugging in the airways (8/8)	Eosinophilia in respiratory tract (8/8) Eosinophilia in gastrointestinal tract (5/8) Pulmonary congestion and/or edema (8/8)	Negative routine toxicological tests (8/8) Serum total IgE level: 318-523 kU/l
Mechanical asphyxia (n = 3)	-	Ligature mark (3/3) Damage to the larynx and associated cartilages (3/3) Damage to the hyoid bone (3/3) Cyanosis (3/3) Conjunctival petechiae (3/3) Congestion and oedema of the face (3/3) Petechiae beneath the lung membrane (3/3)	Alveolar overdistension accompanied by alveolar rupture (3/3) Intra-alveolar hemorrhage with alveolar edema (3/3) Alveolar capillary congestion (3/3)	Negative routine toxicological tests (3/3)
Brain injury (n = 5)	-	Scalp injuries (5/5) Skull fractures (5/5) Intracranial haemorrhage (5/5)	Cerebral contusion (5/5)	Negative routine toxicological tests (5/5)
Acute cardiac death (n = 4)	Previous myocardial infarction (4/4) Hypertension history (3/4)	At least one segment of a major coronary artery must have more than 75% diameter stenosis (4/4) Hemorrhagic halo (4/4) Presence of myocardial scars (4/4)	Coronary atherosclerosis : luminal narrowing reaches over 75% (4/4) Myocardial ischemia (4/4) Myocardial necrosis (4/4) Myocardial granulation tissue (2/4) Myocardial fibrous scars (4/4)	Negative routine toxicological tests (4/4)

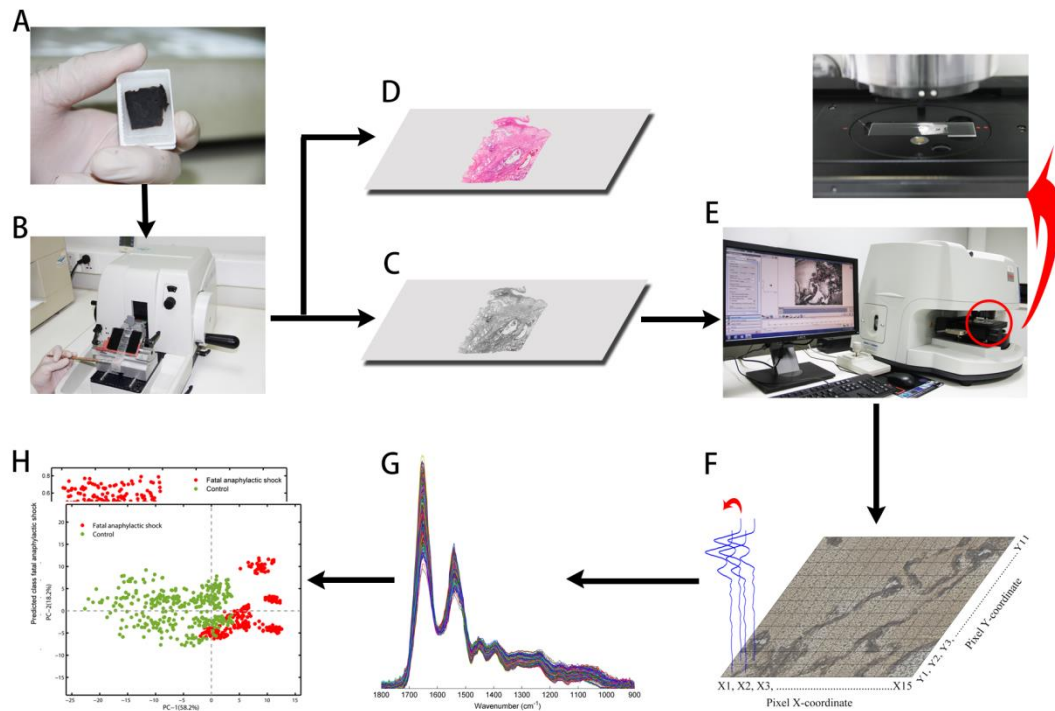


Figure S1 Detail illustration of the experimental workflow.

- 1) **Sample preparation.** For each formalin-fixed paraffin-embedded lung tissue sample (A), 2 adjacent tissue sections were cut using a microtome (B). One section, 7 μm, was deposited on an infrared transparent calcium fluoride (CaF₂) slide (C), and the other, 2.5 μm, was mounted on a conventional glass slide (D). Subsequently, both sections were deparaffinized using standard protocols. After dewaxing, the section transferred onto glass was further stained with hematoxylin/eosin (H&E). The section mounted onto the CaF₂ slide remained stain-free prior to infrared spectral acquisition.
- 2) **Acquisition of spectra (E).**
 - i. Switch on the Nicolet iN10 MX infrared microscope.
 - ii. Fill the detector with liquid N₂.
 - iii. Open the instrument-operating OMNIC Picta software.
 - iv. Apply settings.
 - v. Use the software to get a view of the slide as seen through the microscope.
 - vi. Load the sample onto the stage and focus the microscope.
 - vii. Manipulate the joystick to move the sample around the microscope stage to identify areas that contained abundant pulmonary edema fluid to interrogate.
 - viii. Select a clean, sample-free point on the slide and acquire a background spectrum.
 - ix. Acquire spectra.
- 3) **Spectra selection.** The global biochemical components of each sampled pixel was measured and represented as a spectrum. Spectra from sampled pixels were collected along with coordinates and arranged as “hyperspectral data

cubes” (F). Such a spectral hypercube contained all the pertinent information of the tissue composition and its spatial variations (F). The specific IR spectra of the pulmonary edema fluid were picked out from the “hyperspectral data cubes” and then gathered in a data set (G).

- 4) **Data pre-processing and analysis.** The row spectral data set was preprocessed and then analyzed using pattern recognition methods (H).

Identification of pulmonary edema in forensic autopsy cases of fatal anaphylactic shock using Fourier transform infrared microspectroscopy

Hancheng Lin^{1,2} · Yiwen Luo¹ · Lei Wang¹ · Kaifei Deng¹ · Qiran Sun¹ · Ruoxi Fang² · Xin Wei² · Shuai Zha² · Zhenyuan Wang² · Ping Huang¹

Received: 17 June 2017 / Accepted: 20 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract Anaphylaxis is a rapid allergic reaction that may cause sudden death. Currently, postmortem diagnosis of anaphylactic shock is sometimes difficult and often achieved through exclusion. The aim of our study was to investigate whether Fourier transform infrared (FTIR) microspectroscopy combined with pattern recognition methods would be complementary to traditional methods and provide a more accurate postmortem diagnosis of fatal anaphylactic shock. First, the results of spectral peak area analysis showed that the pulmonary edema fluid of the fatal anaphylactic shock group was richer in protein components than the control group, which included mechanical asphyxia, brain injury, and acute cardiac death. Subsequently, principle component analysis (PCA) was performed and showed that the anaphylactic shock group contained more turn and α -helix protein structures as well as less tyrosine-rich proteins than the control group. Ultimately, a partial least-square discriminant analysis (PLS-DA) model combined with a variables selection method called the genetic algorithm (GA) was built and demonstrated good separation

between these two groups. This pilot study demonstrates that FTIR microspectroscopy has the potential to be an effective aid for postmortem diagnosis of fatal anaphylactic shock.

Keywords Fatal anaphylactic shock · Pulmonary edema fluid · Fourier transform infrared microspectroscopy · Pattern recognition

Introduction

Anaphylaxis is a rapid-onset, multiple-system allergic reaction that can be life-threatening. It often represents an immunological (usually IgE-mediated) or non-immunological reaction that results in degranulation of basophils and mast cells as well as a release of histamine and other inflammatory mediators [1]. According to Yiwen Shen et al., drugs and foods are the two main types of allergens to cause anaphylactic death [2]. In forensic practice, the postmortem diagnosis of anaphylactic shock can be extremely challenging for forensic pathologists due to unclear medical histories, non-specific macroscopic and microscopic findings, and the limited reliability of laboratory tests [3]. Usually, in such cases, the identification of fatal anaphylactic shock is based on exclusion and circumstantial evidence [4–6].

For the past two decades, there has been a focus on investigating biochemical markers that would potentially be suitable for postmortem diagnosis of fatal anaphylactic shock, such as tryptase, histamine, chymase, and IgE [6]. Nevertheless, most of these studies have only focused on plasma fluid. There are few studies focusing on pulmonary edema fluid, which was the most common finding during necropsy of fatal anaphylactic shock cases [2, 7]. Several papers have reported that different types of etiologies are represented in pulmonary edema fluid through different protein contents [8–10].

Hancheng Lin and Yiwen Luo contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00414-017-1721-4>) contains supplementary material, which is available to authorized users.

✉ Zhenyuan Wang
wzy218@xjtu.edu.cn

✉ Ping Huang
huangp@ssfjd.cn

¹ Shanghai Key Laboratory of Forensic Medicine, Shanghai Forensic Service Platform, Institute of Forensic Science, Ministry of Justice, Shanghai 200063, China

² Department of Forensic Pathology, Xi'an Jiaotong University, Xi'an 710061, China

For instance, a study conducted by Hashim et al. [8] indicated that anaphylactic and toxic pulmonary edema fluids were richer in plasma proteins. Alan Fein and coworkers investigated pulmonary edema fluid and plasma proteins in 24 patients involving in neurologic trauma, infection, drug overdose, and hemorrhagic shock [9]. They showed that protein contents of alveolar fluid were various between different etiologies. Ware et al. [10] expanded that study by enlarging the number of patients to 390. Their study results demonstrated the edema fluid-to-plasma protein ratio had excellent diagnostic discrimination between acute lung injury and cardiogenic pulmonary edema. Additionally, Wang et al. [11–13] employed a real-time qPCR technique to investigate the molecular pathology of postmortem pulmonary edema with respect to various causes of death. The results for different mRNA expressions of protein markers indicate the different molecular pathology of alveolar damage arising from different causes of death. All the aforementioned studies taken together imply that the pulmonary edema fluids associated with different etiologies or causes of death contain different biochemical compositions and could be a potential specific biofluid for etiology and cause of death diagnosis. However, to verify that guess, the development of innovative techniques for the further investigation of the biochemical changes of pulmonary edema fluid associated with different etiologies and causes of death is needed.

Fourier transform infrared (FTIR) microspectroscopy is a widely used technique in biomedical research and carried potentials for detecting the global biomolecular composition and variations of biological samples [14]. The greatest benefit of this technique lies in the high molecular sensitivity, which enables the researcher to potentially detect molecular changes that precede any morphological change, thus enabling earlier diagnosis of tissue dysplasia [15]. Currently, the FTIR microspectroscopy technique is widely used for studying cancer diagnosis, such as colon [16], skin [17], cervical [18], breast [19, 20], brain [21], and prostate [22] tumors. Another advantage of this technique is the ability to offer rapid, label-free, and non-destructive measurements of the samples [15]. This is particularly attractive to forensic science evidence, since the researcher could use additional methods to reveal further, complementary information of the samples after analyzing the biochemical content within the samples. To date, this technique has also been used, including the detection of latent fingerprints [23, 24], studying questioned documents [25], bicomponent fibers [26], and trace materials [27] as well as analysis of myocardial infarction [28], diffuse axonal injury [29], and electrical mark [30].

Biofluid-based infrared spectroscopy, an emerging “liquid biopsies” technique for providing specific spectral markers of the biofluids (e.g., plasma, bile, urine, sputum, and cerebrospinal fluid) in the proximity of the diseased cells or tissues, is gaining increasing interest [31]. Several papers have

demonstrated a usefulness of such a technique for diagnosis and monitoring progression of diseases by investigation of plasma [32–38], which contains specific biomarkers, such as glucose, various proteins, creatinine, urea, triglycerides and cholesterol, for various pathological states and diseases. It is reasonable to expect that the pulmonary edema fluid originating mainly from plasma fluid could have a similar biochemical change for the pathological state or disease as the plasma and thus be another specific biofluid for diagnostic purposes.

To the best of our knowledge, FTIR studies on pulmonary edema fluid have not been reported up to now. The main goal of this study was to assess the hypothesis of biochemical differences in edema fluid components between anaphylactic-related death and other causes of death (mechanical asphyxia, brain injury, and acute cardiac death) by means of Fourier transform infrared (FTIR) microspectroscopy. The formalin-fixed paraffin-embedded lung tissue specimens were used for infrared spectroscopy investigation in this study, since such samples are available in large numbers in tissue banks, stable in biochemical composition in long-term conservation, and valuable for retrospective studies. Two pattern recognition methods were introduced in this study. Firstly, principle component analysis (PCA) was performed to analyze variance in protein components between pulmonary edema fluid arising from fatal anaphylactic shock and other causes of death. Subsequently, a partial least square discriminant analysis (PLS-DA) classification model using spectral variables selected by the genetic algorithm (GA) was constructed for distinguishing between anaphylactic death-related pulmonary edema and edema arising from other causes of death.

Material and methods

Sample collection

In this study, 20 formalin-fixed paraffin-embedded lung tissue specimens from 20 forensic autopsy cases were collected from the Department of Forensic Pathology, Institute of Forensic Sciences, Ministry of Justice. The tissue specimens were taken from consistent sites in the lung hilum because this site contains abundant edema fluid. The anaphylaxis shock group comprised 8 cases, and the control group comprised 12 cases (mechanical asphyxia, $n = 3$, brain injury, $n = 5$, acute cardiac death, $n = 4$). Case profiles are given in [Online Resource \(EMS_1\)](#). The diagnoses of the cause of death were determined by professional forensic pathologists through a forensic autopsy examination (including macromorphological, histological, and toxicological examinations), circumstances of death, medical history, and exclusion of other possible cause of death. The main postmortem investigation results pertaining to the cases are given in [Online Resource \(EMS_1\)](#).

Ethical issues

The relevant ethical approval was obtained from the ethics committee of the Institute of Forensic Sciences, Ministry of Justice, P.R. China. All cases selected for this work underwent forensic autopsies requested by a public prosecutor. All tissue specimens were anonymized before analysis.

Sample preparation

From these 20 paraffin-embedded tissue blocks, 7- μm -thick sections were cut using a microtome and transferred onto 1.5-mm-thick infrared transparent calcium fluoride (CaF_2) slides. Then, the samples were dewaxed and dehydrated using xylene and ethanol in that order. An adjacent 2.5- μm section was also cut from each of these tissue blocks, mounted on a glass slide and stained with hematoxylin and eosin (HE) to be used as a morphological reference.

FTIR imaging acquisition and spectral pre-processing

FTIR chemical imaging acquisition was performed in transmission mode using a Nicolet iN10 MX infrared microscope (Thermo Fisher Scientific, Waltham, WA, USA) equipped with a liquid nitrogen-cooled mercury-cadmium-telluride (MCT) detector. Technical detail is given in [Online Resource](#) (EMS_1). Wavenumbers were collected between 4000 and 900 cm^{-1} at a spectral resolution of 4 cm^{-1} . Sixteen scans and 128 scans were conducted for sample and background spectra, respectively. The background spectra were collected on blank areas of CaF_2 slides and automatically subtracted from each sample spectrum. The microscopic aperture was set at $40 \times 40 \mu\text{m}^2$ with a step of 20 μm . The tissue sampling area was nearly $300 \times 220 \mu\text{m}^2$ and a total of 3300 infrared spectra were recorded from 20 images at an average of 165 spectra per image. The spectral collection was performed using OMNIC Picta 9.0 (Thermo Fisher Scientific, Waltham, WA, USA). Row spectral data originating from edema fluid were pre-processed with the Unscrambler 9.7 (Camosoftware, Oslo, Norway). The row spectra were baseline-corrected and then normalized by extending multiplicative signal correction (EMSC) to minimize baseline slope and offset as well as the difference in the thickness of the tissue sections.

Pattern recognition analysis

The analysis of the pre-processed spectral dataset was conducted in Matlab 2014a (Math Works, Natick, MA, USA). Spectral peak area analysis and PCA were performed using the Matlab build-in statistics toolbox. The GA-PLS-DA was performed using classification toolbox 4.2 (Milano

Chemometrics and QSAR Research Group, University of Milano Bicocca, Italy) in combination with the PLS-genetic algorithm toolbox (University of Genoa, Italy).

Pattern recognition is one major area in the chemometrics field [39]. It can be divided into two kinds: unsupervised and supervised. In unsupervised pattern recognition, the data structure is divided into groups by algorithms without providing training standards. By contrast, supervised pattern recognition methods take advantage of information about the class membership of the samples in a certain class to classify new unknown samples.

In this study, one unsupervised (PCA) and one supervised (GA-PLS-DA) pattern recognition method were employed to analyze spectral data with respect to pulmonary edema due to different causes of death. A brief description of these methods is provided below.

PCA is often employed as the first step in data analysis to determine the patterns in the dataset based on differences and similarities between data by calculating the principal component (PC), which is the linear combination of original variables [40]. Usually, the first several PCs explain the majority of the information in the original dataset. The loadings as the coefficients of the corresponding PC are usually used to explain the contribution of variables in the PC. This method is particularly suitable for the spectral dataset for its ability to transform a high-dimension spectral dataset into a low-dimension dataset. In this study, PCA was performed on a spectral dataset containing 293 spectra from edema fluid arising from fatal anaphylactic shock and 345 spectra from the control group. The PC-1 versus PC-2 score plot and corresponding loading plots were graphed to investigate the potential spectral features for distinguishing between the edema fluid of the fatal anaphylactic shock group and the control groups.

PLS-DA is one of the most commonly used supervised techniques for spectroscopic analysis. It is a linear classification method that can decompose X-block with the guide of Y-block and find latent variables (LVs), which are linear combinations of the original variables to maximize the co-variation between X-block and Y-block [41]. In this study, the X-block corresponded to the matrix of edema fluid spectral intensity and the Y-block was represented by the dummy variables of 1 and -1 for the fatal anaphylactic shock and control groups, respectively.

The genetic algorithm is a recent mathematic optimization method inspired by the evolution of species according to the Darwinian theory of the “survival of the fittest” [42]. The application of GA to feature selection of a spectral dataset generally eliminates the irrelevant variables and produces much better results. In this study, GA-PLS-DA was introduced to improve the classification model performance of distinguishing edema fluid arising from fatal anaphylactic shock from fluid associated with three other causes of death.

At the same time, GA-PLS-DA was used to reduce the complexity of the classification model.

Results and discussion

Quality control of pulmonary edema fluid spectra

Each infrared imaging collected from a $300 \times 220 \mu\text{m}^2$ lung tissue section consisted of 15×11 pixels and each pixel represented one spectrum, which reflected the biochemical information of the corresponding area of tissue sections (see Fig. 1a, c). The area selection of spectra for alveolar edema fluid was determined by comparing the results with HE staining (Fig. 1b). Only areas that were unambiguously defined as edema fluid were selected and areas containing or near alveolar walls were not selected (Fig. 1c). Ultimately, 40 unmodified edema fluid spectra were selected randomly from each IR image. Next, the spectral dataset was subjected to a quality test for sample thickness, which discarded spectra with a maximum absorbance less than 0.1 or greater than 1.2. As a result of the test, 638 spectra, including 293 from the fatal anaphylactic shock group and 345 from the control group, were ultimately included. These qualified spectra for alveolar edema fluid were the basis for further analysis.

Spectral peak area analysis

The region from 1800 to 900 cm^{-1} was selected in this study as the bio-fingerprint region of the mid-infrared spectrum [31] (Fig. 2a), which contains the key chemical vibrational bands of biological tissues, such as lipid ester (1800 – 1700 cm^{-1}), amide I and II protein bands (1700 – 1600 cm^{-1} , 1600 – 1500 cm^{-1}), mixed regions containing information from proteins, fatty acids, and phosphate-carrying compounds (1500 – 1200 cm^{-1}) as well as nucleic acids and the carbohydrate main region (1200 – 900 cm^{-1}) [31, 43]. As pulmonary edema fluid is rich in protein, especially plasma protein components, such as albumin and globulin [44], it is significant to investigate amide I and II regions, which are both the most prominent and sensitive vibrational bands of the protein backbone in the mid-infrared spectral region [45–47].

The spectral peak area was often calculated to represent the relative contribution of biochemical components in a spectrum. For this analysis, spectral peak areas of amide I and II regions were calculated on each spectrum (the spectrum was calculated after baseline correction and without EMSC transformation), and the value was then normalized with respect to the spectral peak area of the entire fingerprint region of the corresponding spectrum, 1800 – 900 cm^{-1} . The results are displayed in Fig. 2b, which are expressed as the means $\pm$

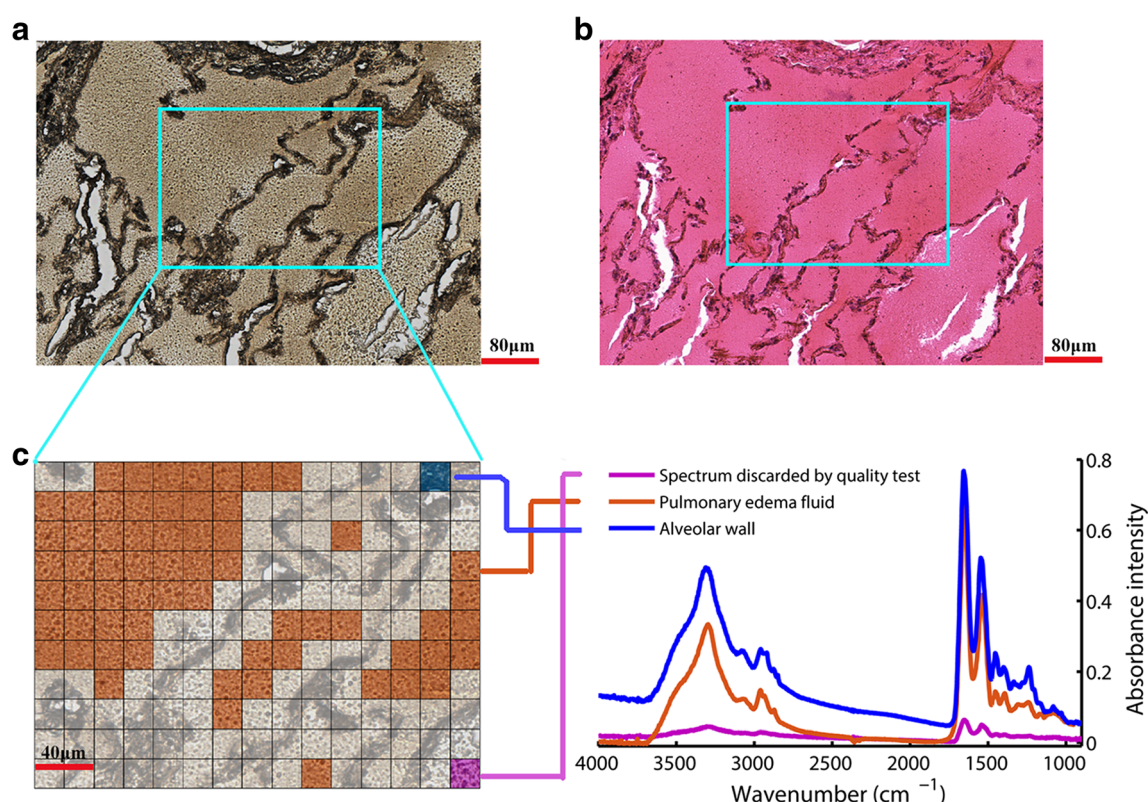


Fig. 1 **a** An unstained white light image of a tissue section of lung hilum from a case of fatal anaphylactic shock with the measured region of interest highlighted in the box. **b** The corresponding reference HE

stained imaging. **c** A detailed view of the measured region containing 15×11 pixels with each pixel representing one spectrum

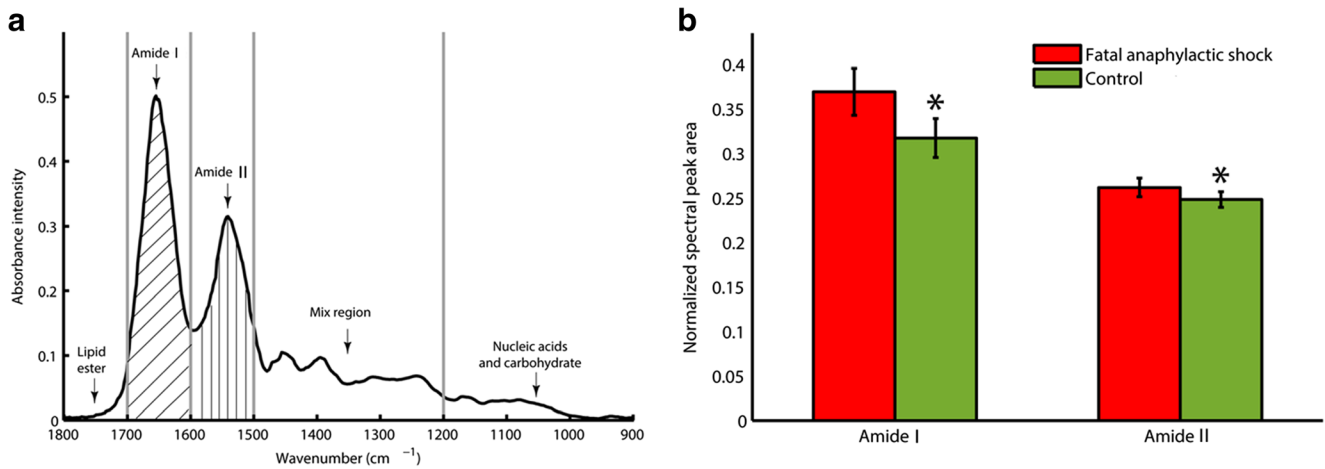


Fig. 2 **a** FTIR average spectrum for pulmonary edema fluid from the fatal anaphylactic shock group in the 1800–900 cm^{-1} region. The spectral peak area of the amide I and II regions was assessed by integration over the ranges of 1700–1600 cm^{-1} and 1600–1500 cm^{-1} . Each value was

then normalized with respect to the spectral peak area of the entire fingerprint spectrum, 1800–900 cm^{-1} . **b** Bar graph of the normalized spectral peak area for amide I and II with respect to the fatal anaphylactic shock and control groups. * $p < 0.05$

standard deviation, and statistical significance was determined using Student's *t* test. As the results show, significantly higher amide I and II levels were found in the fatal anaphylactic shock group compared with the control group, which suggests that the pulmonary edema fluid arising from fatal anaphylactic shock was richer in proteins.

PCA analysis

To further investigate the differences in pulmonary edema fluid protein structures between fatal anaphylactic shock and control groups, PCA was performed on a spectral dataset that was mean-centered and auto-scaled after EMSC transformation with the 1700–1500 cm^{-1} region. Fig. 3a shows the PC-1 versus PC-2 score plot. Although there is little overlapping of scattering points between the fatal anaphylactic shock and control groups, a general discrimination can still be observed along the PC-1, which accounted for 58.2% of total variation and indicated that edema fluid arising from fatal anaphylactic shock

and control groups had a distinct protein composition and the PC-1 retrieved from the original spectral dataset is mainly responsible for that distinction.

Next, a PC-1 loading plot (Fig. 3b) was performed to investigate the basis of the observed separation. Along the direction of PC-1, the discriminating positively correlated loadings for the fatal anaphylactic shock group are mainly found in the 1700–1649 cm^{-1} region, which is a major area for amide I that is abundant in turns and α -helix protein conformational structures [45, 47]. Another feature contributing to the distinction between these two groups was negative loadings in the 1620–1580 cm^{-1} and 1520–1500 cm^{-1} areas, which partly correspond to vibrational modes of side chain residues of tyrosine [48]. The results suggest that edema fluid originating from the fatal anaphylactic shock group contains less tyrosine-rich proteins than that from the control group. Additionally, it is noticeable that the spectra of edema fluid involving fatal anaphylactic shock showed a non-uniform distribution along PC-2. The PC-2 loading plot (Fig. 3c) shows that the main variance is positively correlated with the amide I band at

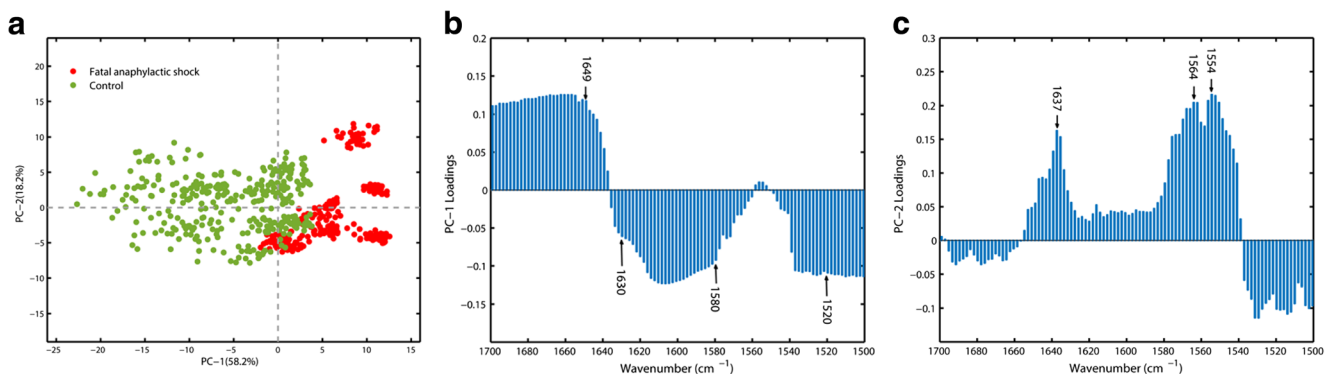


Fig. 3 **a** PCA score plot of EMSC-normalized spectra, including 293 originating from the fatal anaphylactic shock group and 345 from the control group in the 1700–1500 cm^{-1} region. **b** PC-1 correlation loading plot. **c** PC-2 correlation loading plot

1637 cm^{-1} (β -sheet conformation of protein), amide II bands at 1564 cm^{-1} (turn structure of protein), and 1554 cm^{-1} (α -helix structure of protein) [49]. This intra-group conformational differentiation suggests that the protein components of edema fluid in the lung tissues varies among individuals who experience fatal anaphylactic shock, and more samples should be collected for further analysis. Nevertheless, our work demonstrated FTIR spectroscopy can be used to identify protein secondary conformational changes in pulmonary edema fluid, which reflects its high sensitivity to subtle molecular differences in the edema fluid involving different causes of death.

GA-PLS-DA analysis

PCA is as an unsupervised pattern recognition method that has a limited capacity for completely discriminating between fatal anaphylactic shock and control groups. Thus, the next step was to perform PLS-DA since it has been shown that this supervised pattern recognition method is capable of distinguishing different sample categories [50–52].

Prior to discrimination analysis, the genetic algorithm was used to extract significant spectral variables. The parameters used for optimization are as follows: population size, 30 chromosomes; mutation rate, 0.01; crossover rate, 0.5; fitness function, RMSTCV; the number of runs, 100. A total of 252 wavelengths accounting for 54.0% of the full spectral data of 467 wavelengths were ultimately selected and shown in Fig. 4 with blue shades, in which the deeper the blue shade color, the more often the spectral variable covered by the shade was selected in the process of iterations, and the more important

the variable is for distinction. The spectral wavelengths without blue shades indicated that they were eliminated by the GA optimization process. As the result showed, the eliminated spectral variables mainly existed in the lipid ester region and nucleic acids and carbohydrate main region, and the selected spectral variables mainly existed in the amide I and II regions and partly in mixed regions. This result, in combination with PCA results, further demonstrates that spectral varieties of a protein are mainly responsible for the distinction between edema fluid arising from fatal anaphylactic shock and control groups.

Internal cross-validation (CV) is a classic method to select the optimal number of LVs for construction of PLS-DA models [41]. In the process of cross-validation in this study, the spectral dataset was split into ten splits with a Venetian blind approach, nine of which were used for training, while the rest were left out and used for validation. The process was repeated until all splits were predicted once. Ultimately, all ten predicted results were summarized and represented by classification error rate. The lowest possible number of LVs with the lowest CV classification error rate was used to build the models.

The plot of the classification error rate versus the number of LVs is displayed in Fig. 5. Five LVs were ultimately used to build the PLS-DA model. The model was constructed using the extracted dataset containing 638 spectra labeled either fatal anaphylactic shock or control group. The calibrated results showed a distinction between these two groups with an accuracy of 100%. Next, the internal validation was performed through cross-validation to demonstrate the classification

Fig. 4 Selected regions of the infrared spectrum (blue shades) selected by the genetic algorithm

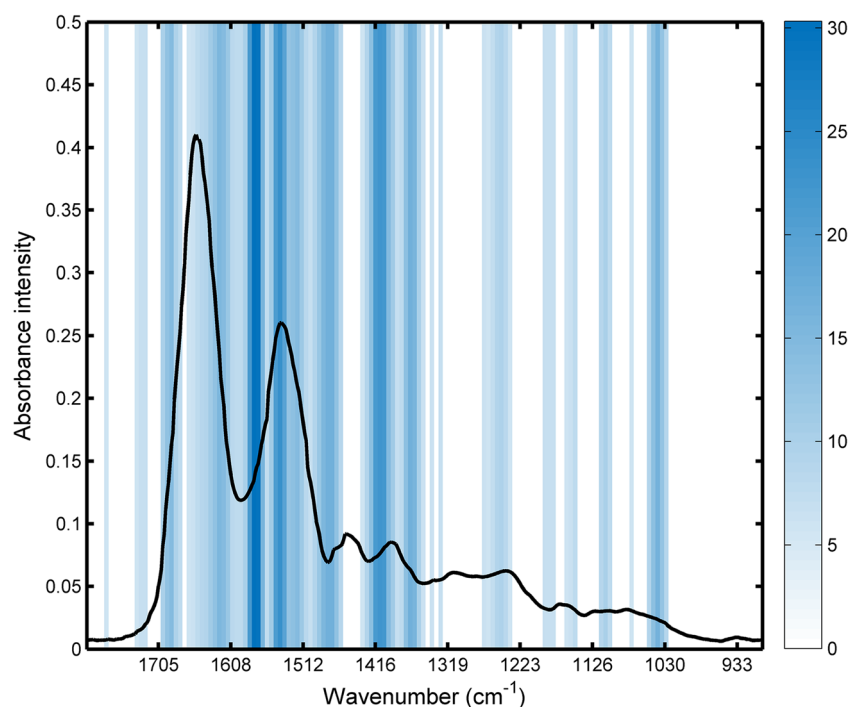
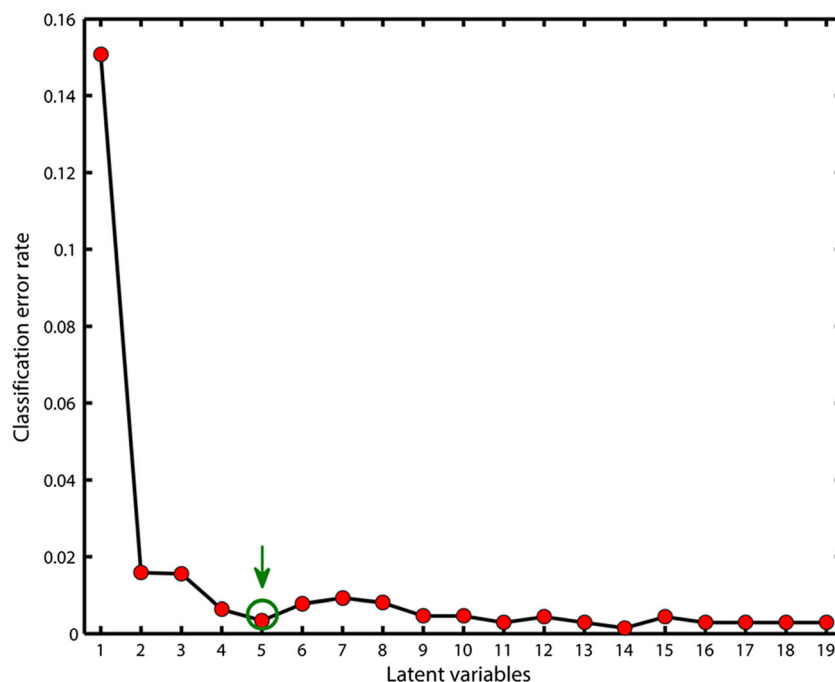


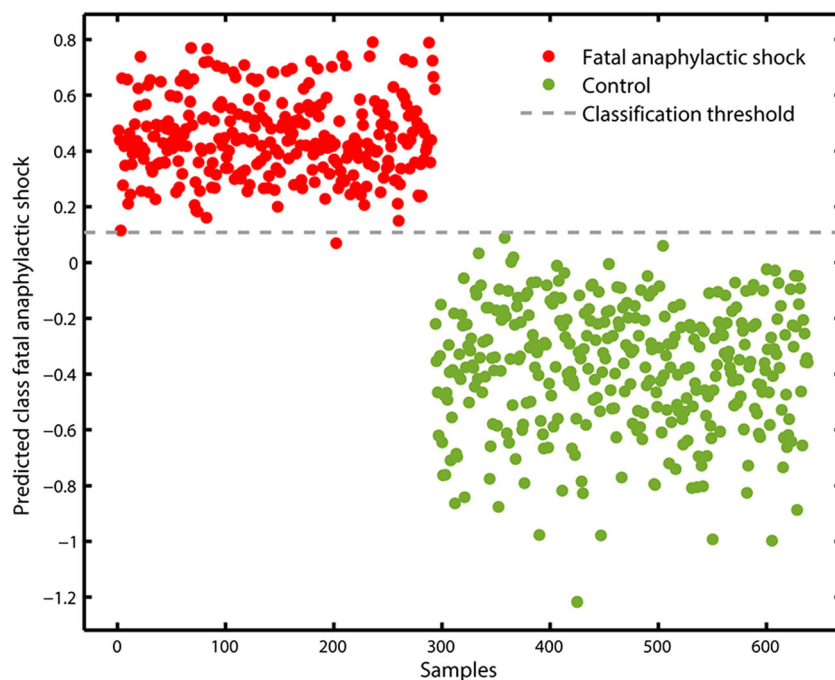
Fig. 5 Classification error rate in cross-validation as a function of the number of latent variables included in the PLS-DA model



ability of PLS-DA model. The validated results are shown in Fig. 6. The internal classification results of the PLS-DA model with respect to fatal anaphylactic shock ability are 0.99 and 1 (sensitivity and specificity, respectively). Only two pulmonary edema fluid spectra originating from the fatal anaphylactic shock group were misclassified into the control group. The rest of the 636 spectra were correctly assigned to their respective classes. These results indicated that this PLS-DA classification model could potentially be stable and reliable.

Additionally, this well-constructed classification model demonstrated a successful extraction of spectral variables by GA, in turn, which further confirmed the significant differences in protein components of pulmonary edema fluid involving fatal anaphylactic shock compared with any other cause of death. According to previous studies, anaphylactic pulmonary edema fluid arises from the high-permeability type of edema, which is characterized by a protein content similar to that of blood [8, 53]. Therefore, it is reasonable to believe

Fig. 6 Cross-validated prediction results for fatal anaphylactic shock class in the binary PLS-DA classification model. The dotted line represents the default classification threshold



that anaphylaxis-related biomarkers or components in plasma may leak into the alveolar edema fluid and alter the vibrational spectroscopic properties, which would result in differences in the infrared spectrum. These particular biomarkers or components were not extensively investigated because they went beyond the capacity of the spectroscopy technique. Nevertheless, one feature of our approach can easily detect these significant differences without needing to know the specific components.

Conclusion

To the best of our knowledge, this is the first study demonstrating that FTIR microspectroscopy combined with pattern recognition methods can be used as an easy and label-free approach to characterize biochemical differences in pulmonary edema fluid arising from different causes of death. Through PCA, the observed spectral variances showed that pulmonary edema fluid from a fatal anaphylactic shock group contained more turn and α -helix protein structures as well as fewer tyrosine-rich proteins than the control group, which included causes of death such as mechanical asphyxia, brain injury, and acute cardiac death. Subsequently, a binary PLS-DA classification model based on spectral variables selected by the genetic algorithm demonstrated good separation between edema fluid spectra from the fatal anaphylactic shock and control groups. These results showed our approach has great potential to complement traditional methods and provide a more accurate postmortem diagnosis of fatal anaphylactic shock. However, before this new method can be employed for forensic practices, much more work needs to be done. For example, it is important to expand the sample sizes for the fatal anaphylactic shock and control groups. It is also necessary to take into consideration additional causes of death in the control group, such as drowning, fatal hyperthermia, fatal hypothermia, and intoxication-related death. According to previous studies, intoxication-related pulmonary edema, which is similar to anaphylactic pulmonary edema, is also a high-permeability type of edema and contains abundant plasma proteins [9]. All of these goals are currently being pursued in our laboratory.

Funding information This project was supported by grants from the National Natural Science Foundation of China (81730056, 81722027, 81471819, 81601645 and 81671869), the National Key R&D Program of China (2016YFC0800701), and the Science and Technology Committee of Shanghai Municipality (17DZ2273200 and 16DZ2290900).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Stone SF, Phillips EJ, Wiese MD, Heddle RJ, Brown SG (2014) Immediate-type hypersensitivity drug reactions. *Br J Clin Pharmacol* 78(1):1–13. <https://doi.org/10.1111/bcp.12297>
2. Shen Y, Li L, Grant J, Rubio A, Zhao Z, Zhang X, Zhou L, Fowler D (2009) Anaphylactic deaths in Maryland (United States) and Shanghai (China): a review of forensic autopsy cases from 2004 to 2006. *Forensic Sci Int* 186(1–3):1–5. <https://doi.org/10.1016/j.forsciint.2008.12.007>
3. Reggiani Bonetti L, Maccio L, Trani N, Radheshi E, Palmiere C (2015) Splenic hypereosinophilia in anaphylaxis-related death: different assessments depending on different types of allergens? *Int J Legal Med* 129(1):97–103. <https://doi.org/10.1007/s00414-014-1004-2>
4. Nishio H, Takai S, Miyazaki M, Horiuchi H, Osawa M, Uemura K, Yoshida K, Mukaida M, Ueno Y, Suzuki K (2005) Usefulness of serum mast cell-specific chymase levels for postmortem diagnosis of anaphylaxis. *Int J Legal Med* 119(6):331–334. <https://doi.org/10.1007/s00414-005-0524-1>
5. Da Broi U, Moreschi C (2011) Post-mortem diagnosis of anaphylaxis: a difficult task in forensic medicine. *Forensic Sci Int* 204(1–3):1–5. <https://doi.org/10.1016/j.forsciint.2010.04.039>
6. Cecchi R (2016) Diagnosis of anaphylactic death in forensics: review and future perspectives. *Legal Med* 22:75–81. <https://doi.org/10.1016/j.legalmed.2016.08.006>
7. Pumphrey RS, Roberts IS (2000) Postmortem findings after fatal anaphylactic reactions. *J Clin Pathol* 53(4):273–276
8. Hashim SW, Kay HR, Hammond GL, Kopf GS, Geha AS (1984) Noncardiogenic pulmonary edema after cardiopulmonary bypass: an anaphylactic reaction to fresh frozen plasma. *Am J Surg* 147(4):560–564
9. Fein A, Grossman RF, Jones JG, Overland E, Pitts L, Murray JF, Staub NC (1979) The value of edema fluid protein measurement in patients with pulmonary edema. *Am J Med* 67(1):32
10. Ware LB, Fremont RD, Bastarache JA, Calfee CS, Matthay MA (2010) Determining the etiology of pulmonary edema by the edema fluid-to-plasma protein ratio. *Eur Respir J* 35(2):331–337
11. Wang Q, Ishikawa T, Michiue T, Zhu BL, Guan DW, Maeda H (2012) Molecular pathology of pulmonary edema after injury in forensic autopsy cases. *Int J Legal Med* 126(6):875–882. <https://doi.org/10.1007/s00414-012-0758-7>
12. Wang Q, Ishikawa T, Michiue T, Zhu BL, Guan DW, Maeda H (2013) Molecular pathology of pulmonary edema in forensic autopsy cases with special regard to fatal hyperthermia and hypothermia. *Forensic Sci Int* 228(1–3):137–141. <https://doi.org/10.1016/j.forsciint.2013.03.007>
13. Du Y, Jin HN, Zhao R, Zhao D, Xue Y, Zhu BL, Guan DW, Xie XL, Wang Q (2016) Molecular pathology of pulmonary edema in forensic autopsy cases with special regard to fatal methamphetamine intoxication. *J Forensic Sci* 61(6):1531–1537. <https://doi.org/10.1111/1556-4029.13199>
14. Trevisan J, Angelov PP, Carmichael PL, Scott AD, Martin FL (2012) Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives. *Analyst* 137(14):3202–3215. <https://doi.org/10.1039/c2an16300d>
15. Kendall C, Isabelle M, Bazanthege mark F, Hutchings J, Orr L, Babrah J, Baker R, Stone N (2009) Vibrational spectroscopy: a clinical tool for cancer diagnostics. *Analyst* 134(6):1029–1045
16. Nallala J, Diebold MD, Gobinet C, Bouche O, Sockalingum GD, Piot O, Manfait M (2014) Infrared spectral histopathology for cancer diagnosis: a novel approach for automated pattern recognition of colon adenocarcinoma. *Analyst* 139(16):4005–4015. <https://doi.org/10.1039/c3an01022h>

17. Ly E, Piot O, Durlach A, Bernard P, Manfait M (2009) Differential diagnosis of cutaneous carcinomas by infrared spectral micro-imaging combined with pattern recognition. *Analyst* 134(6): 1208–1214
18. Walsh MJ, Singh MN, Stringfellow HF, Pollock HM, Hammiche A, Grude O, Fullwood NJ, Pitt MA, Martinhirsch PL, Martin FL (2008) FTIR microspectroscopy coupled with two-class discrimination segregates markers responsible for inter- and intra-category variance in exfoliative cervical cytology. *Biomark Insights* 3(6): 179–189
19. Kumar S, Desmedt C, Larsimont D, Sotiriou C, Goormaghtigh E (2013) Change in the microenvironment of breast cancer studied by FTIR imaging. *Analyst* 138(14):4058–4065. <https://doi.org/10.1039/c3an00241a>
20. Verdonck M, Denayer A, Delvaux B, Garaud S, De Wind R, Desmedt C, Sotiriou C, Willard-Gallo K, Goormaghtigh E (2016) Characterization of human breast cancer tissues by infrared imaging. *Analyst* 141(2):606–619. <https://doi.org/10.1039/c5an01512j>
21. Bergner N, Romeike BF, Reichart R, Kalff R, Krafft C, Popp J (2013) Tumor margin identification and prediction of the primary tumor from brain metastases using FTIR imaging and support vector machines. *Analyst* 138(14):3983–3990. <https://doi.org/10.1039/c3an00326d>
22. Patel IL, Trevisan J, Singh PB, Nicholson CM, Krishnan RK, Matanhelia SS, Martin FL (2011) Segregation of human prostate tissues classified high-risk (UK) versus low-risk (India) for adenocarcinoma using Fourier-transform infrared or Raman microspectroscopy coupled with discriminant analysis. *Anal Bioanal Chem* 401(3):969–982
23. Crane NJ, Bartick EG, Perlman RS, Huffman S (2007) Infrared spectroscopic imaging for noninvasive detection of latent fingerprints. *J Forensic Sci* 52(1):48–53. <https://doi.org/10.1111/j.1556-4029.2006.00330.x>
24. Chen T, Schultz ZD, Levin IW (2009) Infrared spectroscopic imaging of latent fingerprints and associated forensic evidence. *Analyst* 134(9):1902–1904. <https://doi.org/10.1039/b908228j>
25. Dirwono W, Park JS, Agustin-Camacho MR, Kim J, Park HM, Lee Y, Lee KB (2010) Application of micro-attenuated total reflectance FTIR spectroscopy in the forensic study of questioned documents involving red seal inks. *Forensic Sci Int* 199(1–3):6–8. <https://doi.org/10.1016/j.forsciint.2010.02.009>
26. Flynn K, O'Leary R, Roux C, Reedy BJ (2006) Forensic analysis of bicomponent fibers using infrared chemical imaging. *J Forensic Sci* 51(3):586–596. <https://doi.org/10.1111/j.1556-4029.2006.00116.x>
27. Chan KL, Kazarian SG (2006) Detection of trace materials with Fourier transform infrared spectroscopy using a multi-channel detector. *Analyst* 131(1):126–131. <https://doi.org/10.1039/b511243e>
28. Yang TT, Weng SF, Zheng N, Pan QH, Cao HL, Liu L, Zhang HD, Mu d W (2011) Histopathology mapping of biochemical changes in myocardial infarction by Fourier transform infrared spectral imaging. *Forensic Sci Int* 207(1–3):e34–e39. <https://doi.org/10.1016/j.forsciint.2010.12.005>
29. Yang T, He G, Zhang X, Chang L, Zhang H, Ripple MG, Fowler DR, Li L (2014) Preliminary study on diffuse axonal injury by Fourier transform infrared spectroscopy histopathology imaging. *J Forensic Sci* 59(1):231–235. <https://doi.org/10.1111/1556-4029.12290>
30. Zhang J, Lin W, Lin H, Wang Z, Dong H (2017) Identification of skin electrical injury using infrared imaging: a possible complementary tool for histological examination. *PLoS One* 12(1):e0170844. <https://doi.org/10.1371/journal.pone.0170844>
31. Baker MJ, Hussain SR, Lovergne L, Untereiner V, Hughes C, Lukaszewski RA, Thiefin G, Sockalingum GD (2016) Developing and understanding biofluid vibrational spectroscopy: a critical review. *Chem Soc Rev* 45(7):1803–1818. <https://doi.org/10.1039/c5cs00585j>
32. Erukhimovitch V, Talyshinsky M, Souprun Y, Huleihel M (2006) FTIR spectroscopy examination of leukemia patients plasma. *Vib Spectrosc* 40(1):40–46
33. Ahmed SSSJ, Santosh W, Kumar S, Christlet THT (2010) Neural network algorithm for the early detection of Parkinson's disease from blood plasma by FTIR micro-spectroscopy. *Vib Spectrosc* 53(2):181–188
34. Gajjar K, Trevisan J, Owens G, Keating PJ, Wood NJ, Stringfellow HF, Martinhirsch PL, Martin FL (2013) Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. *Analyst* 138(14):3917
35. Staniszewskaslezak E, Fedorowicz A, Kramkowski K, Leszczynska A, Chlopicki S, Baranska M, Malek K (2015) Plasma biomarkers of pulmonary hypertension identified by Fourier transform infrared spectroscopy and principal component analysis. *Analyst* 140(7):2273–2279
36. Zelig U, Barlev E, Bar O, Gross I, Flomen F, Mordechai S, Kapelushnik J, Nathan I, Kashtan H, Wasserberg N (2015) Early detection of breast cancer using total biochemical analysis of peripheral blood components: a preliminary study. *BMC Cancer* 15(1):408
37. Staniszewskaslezak E, Mateuszuk L, Chlopicki S, Baranska M, Malek K (2016) Alterations in plasma biochemical composition in NO deficiency induced by L-NAME in mice analysed by Fourier Transform Infrared Spectroscopy. *J Biophotonics* 9(10): 1098
38. Staniszewskaslezak E, Wiercigroch E, Fedorowicz A, Buczek E, Mateuszuk L, Baranska M, Chlopicki S, Malek K (2017) A possible FTIR-based plasma fingerprint of ACE-I induced reversal of endothelial dysfunction in diabetic mice. *J Biophotonics*. <https://doi.org/10.1002/jbio.201700044>
39. Barry L, Workman JJ (2004) Chemometrics. *Anal Chem* 76(12): 3365–3371
40. Bro R, Smilde AK (2014) Principal component analysis. *Anal Methods* 6(9):2812. <https://doi.org/10.1039/c3ay41907j>
41. Ballabio D, Consonni V (2013) Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal Methods* 5(16):3790. <https://doi.org/10.1039/c3ay40582f>
42. Leardi R, González AL (1998) Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemom Intell Lab Syst* 41(2):195–207
43. Helm D, Labischinski H, Schallehn G, Naumann D (1991) Classification and identification of bacteria by Fourier-transform infrared spectroscopy. *J Gen Microbiol* 137(1):69
44. Kobayashi T, Nitta K, Ganzuka M, Inui S, Grossmann G, Robertson B (1991) Inactivation of exogenous surfactant by pulmonary edema fluid. *Pediatr Res* 29(4):353–356
45. Barth A (2007) Infrared spectroscopy of proteins. *Biochim Biophys Acta* 1767(9):1073–1101. <https://doi.org/10.1016/j.bbabi.2007.06.004>
46. Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, Fielden PR, Fogarty SW, Fullwood NJ, Heys KA, Hughes C, Lasch P, Martin-Hirsch PL, Obinaju B, Sockalingum GD, Sule-Suso J, Strong RJ, Walsh MJ, Wood BR, Gardner P, Martin FL (2014) Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* 9(8):1771–1791. <https://doi.org/10.1038/nprot.2014.110>
47. Yang H, Yang S, Kong J, Dong A, Yu S (2015) Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nat Protoc* 10(3):382–396. <https://doi.org/10.1038/nprot.2015.024>
48. Gabashvili IS, Menikh A, Ségui J, Fragata M (1998) Protein structure of photosystem II studied by FT-IR spectroscopy. Effect of digalactosyldiacylglycerol on the tyrosine side chain residues. *J Mol Struct* 444(1–3):123–133

49. Darwish SM (2010) Spectroscopic study of propofol binding to human serum albumin. *Biophys Rev Lett* 05(04):209–226. <https://doi.org/10.1142/s1793048010001202>
50. de Almeida MR, Correa DN, Rocha WFC, Scafi FJO, Poppi RJ (2013) Discrimination between authentic and counterfeit banknotes using Raman spectroscopy and PLS-DA with uncertainty estimation. *Microchem J* 109:170–177. <https://doi.org/10.1016/j.microc.2012.03.006>
51. Zhang J, Li B, Wang Q, Li C, Zhang Y, Lin H, Wang Z (2017) Characterization of postmortem biochemical changes in rabbit plasma using ATR-FTIR combined with chemometrics: a preliminary study. *Spectrochim Acta A* 173:733–739. <https://doi.org/10.1016/j.saa.2016.10.041>
52. Mistek E, Lednev IK (2015) Identification of species' blood by attenuated total reflection (ATR) Fourier transform infrared (FT-IR) spectroscopy. *Anal Bioanal Chem* 407(24):7435–7442. <https://doi.org/10.1007/s00216-015-8909-6>
53. Jr JBS, Hattori T, Ray CS, Bove AA, Cianci P (2001) Pulmonary edema associated with scuba diving : case reports and review. *Chest* 120(5):1686–1694