



OPEN

Micro-Fourier-transform infrared reflectance spectroscopy as tool for probing IgG glycosylation in COVID-19 patients

Carla Carolina Silva Bandeira¹, Karen Cristina Rolim Madureira², Meire Bocoli Rossi², Juliana Failde Gallo², Ana Paula Marques Aguirra da Silva², Vilanilse Lopes Torres², Vinicius Alves de Lima³, Norival Kesper Júnior³, Janete Dias Almeida⁴, Rodrigo Melim Zerbinati³, Paulo Henrique Braz-Silva^{3,5}, José Angelo Lauletta Lindoso^{2,3,6} & Herculano da Silva Martinho¹✉

It has been reported that patients diagnosed with COVID-19 become critically ill primarily around the time of activation of the adaptive immune response. However the role of antibodies in the worsening of disease is not obvious. Higher titers of anti-spike immunoglobulin IgG1 associated with low fucosylation of the antibody Fc tail have been associated to excessive inflammatory response. In contrast it has been also reported that NP-, S-, RBD- specific IgA, IgG, and IgM are not associated with SARS-CoV-2 viral load, indicating that there is no obvious correlation between antibody response and viral antigen detection. In the present work the micro-Fourier-transform infrared reflectance spectroscopy (micro-FTIR) was employed to investigate blood serum samples of healthy and COVID-19-ill (mild or oligosymptomatic) individuals (82 healthcare workers volunteers in “Instituto de Infectologia Emilio Ribas”, São Paulo, Brazil). The molecular-level-sensitive, multiplexing quantitative and qualitative FTIR data probed on 1 μL of dried biofluid was compared to signal-to-cutoff index of chemiluminescent immunoassays CLIA and ELISA (IgG antibodies against SARS-CoV-2). Our main result indicated that 1702–1785 cm^{-1} spectral window (carbonyl C=O vibration) is a spectral marker of the degree of IgG glycosylation, allowing to probe distinctive sub-populations of COVID-19 patients, depending on their degree of severity. The specificity was 87.5 % while the detection rate of true positive was 100%. The computed area under the receiver operating curve was equivalent to CLIA, ELISA and other ATR-FTIR methods (> 0.85). In summary, overall discrimination of healthy and COVID-19 individuals and severity prediction as well could be potentially implemented using micro-FTIR reflectance spectroscopy on blood serum samples. Considering the minimal and reagent-free sample preparation procedures combined to fast (few minutes) outcome of FTIR we can state that this technology is suitable for fast screening of immune response of individuals with COVID-19. It would be an important tool in prospective studies, helping investigate the physiology of the asymptomatic, oligosymptomatic, or severe individuals and measure the extension of infection dissemination in patients.

Accumulating evidence suggests that a subset of patients with severe COVID-19 may have cytokine storm syndrome¹. It has been recommended that hyperinflammation should be recognized and treated using already approved therapies with proven safety profiles to reduce the increasing mortality¹.

¹Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP 09210-580, Brazil. ²Instituto de Infectologia Emilio Ribas, São Paulo, Sp 01246-900, Brazil. ³Instituto de Medicina Tropical de São Paulo, Universidade de São Paulo, São Paulo, SP 05403-000, Brazil. ⁴Departamento de Biociências e Diagnóstico, Instituto de Ciência e Tecnologia, Universidade Estadual Paulista, São José dos Campos, SP 12245-000, Brazil. ⁵Faculdade de Odontologia Departamento de Estomatologia, Universidade de São Paulo, São Paulo, SP 05508-000, Brazil. ⁶Departamento de Moléstias Infecciosas e Parasitárias, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP 01255-090, Brazil. ✉email: herculano.martinho@ufabc.edu.br

It has been noticed that patients diagnosed COVID-19 become severely ill, especially at the time of activation of the adaptive immune response^{2–4}. Hoepel et al.² reported that antibodies play a role in disease exacerbation at the time of seroconversion. They found that the excessive inflammatory response depends on two antibody features which are specific to patients with severe COVID-19: (1) higher titers of anti-spike IgG and (2) the anti-spike IgG of patients with severe COVID-19 is inherently more pro-inflammatory due to different glycosylation, particularly low fucosylation, of the Fc tail of the antibody. Analogously Chakraborty et al.³ suggested that patients with severe COVID-19 have a unique serological signature, including an increased likelihood of IgG1 with afucosylated Fc glycans. On the other hand, Luo et al.⁴ showed that NP-, S-, RBD-specific IgA, IgG, and IgM were not associated with SARS-CoV-2 viral load, suggesting that there is no obvious correlation between antibody response and viral antigen detected in nasopharyngeal swabs. They analyzed antibody and cytokine responses in COVID-19 from asymptomatic to severe patients (123 serum samples from 63 COVID-19 patients) and evaluated the impact of various risk factors, including comorbidities, male sex, and advancing age on the host immune response COVID-19 patients. These antagonistic data highlight the relevance of investigating structural aspects of serum IgG from COVID-19 patients to establish its role as a severity marker.

Vibrational spectroscopic techniques such as Fourier-transform infrared absorption (FTIR) have been successfully used to study biological samples. They provide quantitative and qualitative multiplexed information at the molecular level, even when triggered by subtle changes in a sample⁵. Vibrational spectroscopy involves label-free techniques that allow detection of electronic changes in the internal vibrational energy levels of biomolecules. It has been used in the study of cells, tissues and biofluids and provided valuable insights into pathologies^{5–8}. Barauna et al.⁹ investigated saliva samples from throat swabs using Attenuated Total Reflectance FTIR (ATR-FTIR). Initial classification of swab samples into negative and positive COVID-19 infections was based on symptoms and PCR testing results ($n = 111$ negatives and 70 positives). A blind sensitivity of 95% and specificity of 89% were achieved. Artificial saliva samples containing inactivated γ -irradiated COVID-19 virus particles at concentrations up to 1582 copies/mL were also analyzed to understand the spectral response of the virus⁹. Proof-of-concept study by Nogueira et al.¹⁰ conducted on 243 patients indicated that ATR-FTIR could be a cost-effective solution for high-throughput screening of suspect patients for COVID-19 using oropharyngeal swab suspension fluid. Sensitivity, specificity, and accuracy better than 84%, 64%, and 76.9% were found, respectively. Kitane et al.¹¹ reported a method for COVID-19 discrimination based on spectral ATR-FTIR analysis of 280 RNA extracts from nasopharyngeal samples (100 SARS-CoV-2 PCR-positive patients and 180 SARS-CoV-2 PCR-negative patients). The proposed method is based on ATR-FTIR analysis of the extracted RNA and machine learning modeling. The authors report that this technique achieves 97.8% accuracy, 97% sensitivity, and 98.3% specificity while reducing the testing time from hours to minutes after RNA extraction. The reported area under the ROC curve (*AUC*) ranged from 0.54 to 1, depending on the statistical model used. Wood et al.¹² proposed a portable infrared spectrometer with custom-made transfectance accessories for rapid point-of-care detection of COVID-19 markers in saliva. They tested the system on samples from 29 subjects who tested positive for SARS-CoV-2 by RT-PCR and 28 who were negative, and achieved a sensitivity of 93% and a specificity of 82%. However the work of Wood et al. did not aim to explore the rich biochemical information contained in the spectral data.

Some clues concerning immune response and COVID-19 had been considered by Dogan et al.¹³. They used ATR-FTIR to investigate serum samples from 56 patients and revealed that the CoronaVac-Sinovac COVID-19 vaccine administration induced significant changes in some functional groups belonging to lipids, proteins and nucleic acids. FTIR has been used to investigate the secondary structural composition and changes in structural dynamics of IgG upon glycation, oxidation and glycooxidation¹⁴. Thus, it is a suitable technology to study structural changes in IgG induced by COVID-19.

Aiming to investigate possible correlation between antibody response and SARS-CoV-2 infection blood serum samples from adult healthcare workers with COVID-19 (asymptomatic or mild level) and healthy individuals were investigated here by micro-FTIR reflectance spectroscopy. Detailed characterization of the immune response of asymptomatic or oligosymptomatic individuals is important for prospective studies to describe the physiology of these cases compared to cases with higher severity and also to determine the extent of spread of infection by these patients.

Methods

Sampled population. Serum samples from 82 healthcare workers volunteers of the “Instituto de Infectologia Emilio Ribas”, São Paulo, Brazil were included in this study. These samples are part of a cohort study to evaluate the seroprevalence of SARS-CoV-2 infection in healthcare population. Samples were collected from July to November of 2020, before the beginning of the vaccination program in Brazil. Levels of illness were mild or oligosymptomatic. This study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by Research Ethics Committee of the “Instituto de Infectologia Emilio Ribas”, São Paulo, Brazil, protocol number CAAE 32264120.5.2001.0061. Invited volunteers were informed about the objectives, propositions and conditions of this project, and those who agreed to participate in the research signed the free and informed consent term. A volume of 5 mL of whole blood was collected by peripheral vein puncture and stored in a 10 mL dry tube for each patient. Then it was centrifuged at 2052 g for 10 min to separate the serum. Two mL of serum were stored in a -20°C freezer until performing the immunoassays and FTIR tests.

Detection of antibodies anti-SARS-CoV-2. *Chemiluminescent immunoassay (CLIA).* A CLIA (IgG Antibodies against SARS-CoV-2; reagents pack # 619 9919; VITROS Immunodiagnostic calibrator # 619 992; Ortho Clinical Diagnostics, Raritan, NJ, USA) immunoassay was performed to detect immunoglobulin G anti-Spike protein from SARS-CoV-2. This assay does not differentiate binding IgG antibodies from virus-neutralizing IgG antibodies¹⁵. The results were expressed in terms of ratio of the sample signal to a calibrator-assigned

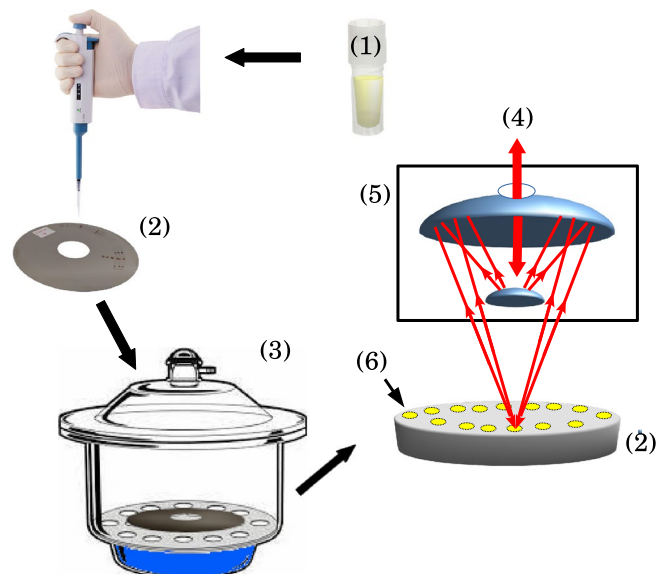


Figure 1. Scheme for micro-FTIR reflectance measurements for human serum. One μL of diluted (1 : 3 in ultra-pure water) serum sample solution (1) was transferred to a circular platinum sample holder (2). Then the sample holder was installed in a desiccator with saturated solution of NaCl (3) which controls the relative moisture in 80%. The drying time was 10 min. After this period the sample holder was installed on the FTIR reflectance accessory of the micro-FTIR spectrometer. The IR beam (4) passing through the IR $60\times$ magnification Cassegrain lens (5) focuses the light on a given sample (6). The reflected light is collected by the same lens and analyzed by the spectrometer.

cutoff signal with threshold of 1.0. Negative diagnosis were considered for outcomes with reactivity index < 1 otherwise they were considered positive¹⁶. The test sensitivity was reported to be 90.0% (≥ 8 days) and specificity of 100.0%¹⁶. This CLIA diagnostic test uses S antigens for SARS-CoV-2 detection¹⁷.

Enzyme-linked immunosorbent assay (ELISA). Anti-SARS-CoV-2 IgG ELISA (Euroimmun Medizinische Labordiagnostika, Lübeck, Germany; # EI 2606-9601 G; Indirect ELISA) assay was performed to detect IgG antibodies against SARS-CoV-2. In this diagnostic test IgG antibodies against SARS-CoV-2 spike protein subunit 1 (S1) are detected in human serum or plasma. Following the instructions of the manufactures, serum samples were 1 : 101 diluted, added to wells coated with recombinant SARS-CoV-2 antigen and incubated for 60 min at 37°C . Then, wells were washed three times and followed by the addition of HRP-conjugated anti-human IgG and subsequent incubation for 30 min at 37°C . Wells were washed three times and a chromogen solution was added. After 30 min of incubation at room temperature, the reaction was stopped and the absorbance at 450 nm with reference at 620 nm was read on a microplate reader. A ratio between the extinction of the sample and calibrator on each plate were calculated. According to the recommendations of the manufacturer, a signal-to-cutoff ratio smaller than 0.8 is considered negative, while a positive one if greater than 1.1. The borderline region falls into 0.8–1.1 interval. The sensitivity and specificity of this assay were reported to be 90.0% and 100%, respectively¹⁷.

Fourier-transform infrared absorption (FTIR) reflectance spectroscopy. All samples were brought to room temperature prior to preparation for micro-FTIR measurements. Aliquots of $1\ \mu\text{L}$ of serum (1 : 3 in ultrapure water) were deposited in platinum sample holder and dried at 80% relative moisture under a desiccator with NaCl saturated solution in order to avoid coffee ring effect and obtain a homogeneous biofilm¹⁸. The final droplet presented an average radius of $1000 \pm 150\ \mu\text{m}$. A Varian 610 FT-IR micro-spectrometer equipped with a linearized MCT detector with a $100 \times 100\ \mu\text{m}$ detector element (InfraRed Associates, Inc.) coupled to a 640-IR FT-IR spectrometer was used in reflectance mode for spectra acquisition. After optical focusing, the microscope aperture in the Cassegrain collecting lens ($60\times$ magnification) was reduced to $150 \times 150\ \mu\text{m}^2$ to set the area of interest for measurement on the sample. The spectral resolution was set to $2\ \text{cm}^{-1}$ in order to obtain the maximum throughput in the IR microspectrometer. The number of scans was 32 per sample (equivalent to 30 s time of acquisition). The scheme of methodology is shown on Fig. 1.

Statistical analysis. *Principal components analysis (PCA).* The classical Principal Components Analysis (PCA)¹⁹ was performed on mean centered raw data to extract outliers and identify possible experimental bias. All spectral analysis steps were implemented in the ChemSpec vignette available in the software R²⁰. Outliers were identified using the Q and T^2 Hotelling's statistics. The Q statistics indicates how well each observation matches to the PCA model and the Q -residuals measure the residual between a sample and its projection on the factors retained in the model. Large residual outliers can be detected by inspection of Q -residuals. On the other

Variable	Positive ($n = 33$)	Negative ($n = 49$)	p -value
Female	76.7%	72.1%	< 5%
Male	23.3%	27.9%	< 5%
Age	44 ± 14	49 ± 12	< 5%
Comorbidities	39.4%	16.3%	45.4%
Body mass index (kg/m ²)	25 ± 12	30 ± 12	< 5%

Table 1. Demographic data of volunteers, a set of 82 healthcare workers from “Instituto de Infectologia Emilio Ribas”, São Paulo, Brazil. The list of comorbidities includes rheumatoid arthritis, asthma, diabetes, systemic arterial hypertension, chronic obstructive pulmonary disease, obesity, and hypothyroidism.

hand, Hotelling's T^2 value represents a measure of the variation in each sample within the model, indicating how far each sample is from the center (scores = 0) of the model. It is a quantifier for scores outliers. The T^2 Hotelling's versus Q -residuals (reduced) plot were inspected raw spectral data.

Partial least squares: discriminant analysis (PLS-DA). All spectra were pre-processed to become comparable for statistical analysis. The baseline was corrected using the least-squares polynomial curve fitting method as described by Lieber and Mahadevan-Jansen²¹. All spectra were normalized and scaled using probabilistic quotient normalization²². Then PLS-DA analysis was performed. It is a multivariate supervised method that uses linear regression of original variables to predict the class membership. In our case the PLS regression was performed using the *pls* function provided by R *pls* package^{20,23}. The classification and cross-validation were performed using the corresponding wrapper function using the *caret* package²⁰. A permutation test was performed to assess the performance of class discrimination. In each permutation, a PLS-DA model was built between the data and the permuted class labels using the optimal number of components determined by leave-one-out cross validation for the model based on the original class assignment. The class discrimination performance was measured using classification accuracy, R^2 , and Q^2 parameters. The first one is based on prediction accuracy. The R^2 parameter is the “goodness of fit” or explained variation which is based on the ratio of the between group sum of the squares and the within group sum of squares. On the other hand, Q^2 is the “goodness of prediction”, or predicted variation, calculated from cross validation. In each round, the predicted data are compared with the original data, and the sum of squared errors is calculated and then summed over all samples (Predicted Residual Sum of Squares or PRESS). For convenience, the PRESS is divided by the initial sum of squares and subtracted by 1 to scale to R^2 . Good predictions will have low PRESS or high Q^2 while negative Q^2 means the model is not predictive at all or is overfitted^{8,24,25}. Two quantifiers were used to measure the vibrational band frequency importance in PLS-DA model. The first, Variable Importance in Projection (VIP) is a weighted sum of squares of the PLS loadings taking into account the amount of explained spectral intensity-variation in each dimension. The other importance measure is based on the weighted sum of PLS regression. The weights are a function of the reduction of the sums of squares across the number of PLS components. For multiple-group analysis, the same number of predictors will be built for each group and the average of the feature coefficients were used to indicate the overall coefficient-based importance. The receiver operating characteristic (ROC) analysis was used to evaluate the discriminating performance and the area under the ROC curve (AUC) as its summary index. In general tests with excellent discriminating capability will furnish $AUC > 0.80$ ^{26,27}.

Normality and F tests. The Kolmogorov-Smirnov test of normality²⁸ was applied to test the hypothesis that IgG data does not differ significantly from that which is normally distributed. The F-test was used to test the hypothesis of equality of averages of two sets of data populations of unequal size (IgG and demographic data)²⁹.

Results and discussion

The main clinical and demographic pieces of information about volunteers are presented on Table 1. Percentages of 40.2% ($n = 33$) and 59.8% ($n = 49$) of individuals tested positive and negative, respectively. Female individuals predominated in the sampled groups (> 72 % of total number of patients). However we notice that this unbalance did not induce a bias on the positive or negative rates within the confidence value of 5%. The average age of patients in positive and negative groups was 44 ± 14 years and 49 ± 12 years, respectively. Again these differences were not statistically relevant for COVID-19 diagnostic purposes. Likewise the body mass index did not represent a statistically relevant variable comparing negative and positive groups. On the other hand, the presence of comorbidities presented a well-defined increased risk for positive test against COVID-19. A percentage of 39.4% of individuals which tested positive presented some kind of comorbidity (rheumatoid arthritis, asthma, diabetes, systemic arterial hypertension, chronic obstructive pulmonary disease, obesity, or hypothyroidism) while this rate is only 16.3% in negative group. The outcome of F-test indicated that there is enough evidence against the hypothesis that the population sampled with positive and negative test has the same average ($p = 45.4\%$). The correlation between comorbidities and prevalence of COVID-19 is well reported in literature and our findings give one more piece of evidence concerning this important aspect of etiology of COVID-19³⁰. We notice that CLIA and ELISA tests were 100% concordant about the patient's diagnosis in our cohort.

Average FTIR spectra for negative (black line) and positive (red line) classes of samples in the fingerprint spectral window (880–1800 cm^{-1}) are shown on Fig. 2. Assignments for the main vibrational bands (vertical

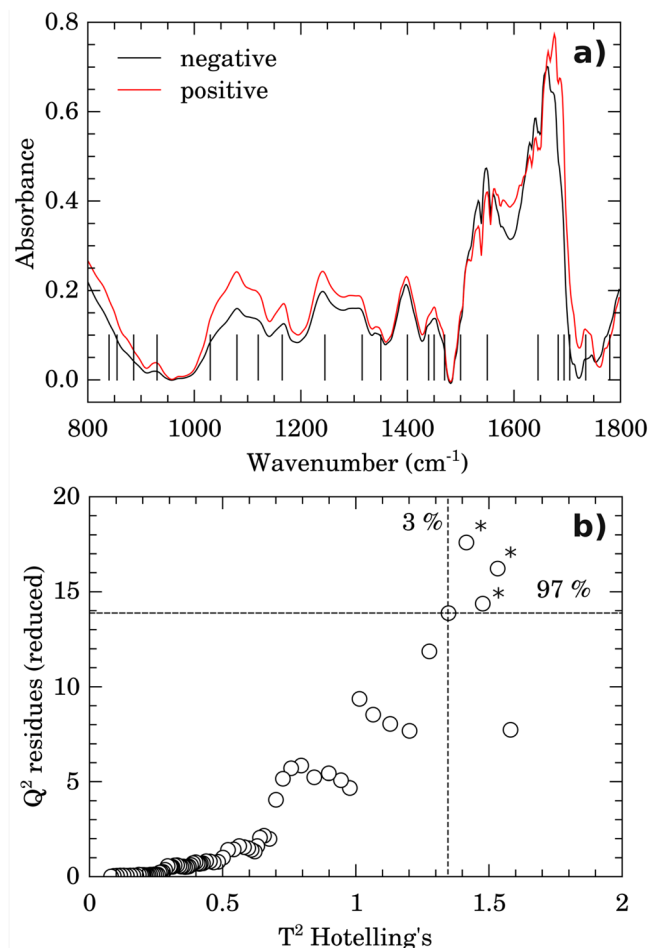


Figure 2. Average spectra and outliers. **(a)** Average micro-FTIR blood serum spectra for negative (black line) and positive (red line) groups. The vertical lines represent the main vibrational bands contributing to discrimination of groups (see band assignments on Table 2). **(b)** Outliers identification by inspection of Q^2 residuals (reduced) versus T^2 Hotelling's. The outliers were indicated by “*”. Dashed horizontal lines and vertical lines represent confidence limits of 3% (Hotelling's T^2) and 97% (Q residuals), respectively.

lines in Fig. 2a) are presented on Table 2. Proteins, amino acids, and nucleic acids vibrational bands dominate the spectra in accordance with the average chemical composition of the blood serum³¹.

Prior data processing and multivariate statistical analysis, a quality check evaluation was performed on raw spectral data to identify anomalous spectra, outliers and/or biased patterns. The PCA was calculated on mean-centered raw data and then Q residuals (reduced) versus T^2 -Hotelling's plot^{8,25} was checked (Fig. 2b) in order to find residuals and scores outliers. Data outside of the confidence limits of 97% for scores and 3% for residuals were considered outliers (indicated by “*” in Fig. 2b) and removed in further analysis. PLS-DA analysis was then performed on processed, normalized, and scaled spectra. The optimal number of factors was determined by cross-validation after inspection of accuracy, R^2 , and Q^2 . Figure 3 presents the pairwise scores up to 5th PC (4.7% of explained variance) for positive and negative groups. At first glance, combinations including components 1 and 2 are prone to discriminate between positive and negative samples.

The best performance of PLS-DA classification was observed considering 2 components (Fig. 4a). In this case the observed accuracy on groups discrimination was 76% while $R^2 = 0.39$ and $Q^2 = 0.34$ (Fig. 4a). Regression coefficients are shown in Fig. 4b). Both curves for discriminating positive and negative groups appeared smooth, showing no random fluctuations around positive and negative values, which would be a symptom of overfitting. However, the coefficients for positive and negative classes appeared superimposed in many spectral windows, indicating greater similarity among them. The calculated response is shown in Fig. 4d). Sensibility and specificity were 53.1% and 87.5%, respectively in this case. Interestingly, there is a distinct subgroup of misclassified samples (labeled here as “mix” group). This subgroup showed a clear spectral signature, as seen in the heatmap in Fig. 5.

The performance of the PLS-DA model with 3 possible groups (positive, negative, and mix) is shown in Fig. 4c). R^2 and Q^2 presented a consistent increase of ~ 0.1 while the accuracy was around 76%. The coefficient and response obtained are indicated in green in Fig. 4b,d, respectively. Specificity was 87.5% while the detection rate of true positive and true mix increased to 100%. However, one outstanding result was observed contrasting the IgG Signal-to-Cutoff index from CLIA of each positive, mix, and negative groups to FTIR PLS-DA

Wavenumber (cm ⁻¹)	Assignment	References
840–845	Left-handed helix DNA (Z form)	32
850–856	C _{2'} endo/anti of deoxyribose in B-form helix conformation	32
886	C–C, C–O deoxyribose	32
930	Left-handed helix DNA (Z form)	32
1030	Stretching C–O ribose	32
1080	Ring stretching vibrations in phenylalanine, tryptophan or tyrosine	32
1120	Symmetric stretching P–O–C, phosphorylated saccharide residue	32
1165	C–O stretching mode of C–OH groups of serine, threonine, tyrosine	32
1245	Amide III α -helix conformation of proteins	32
1315	Amide III of proteins	32
1350	Stretching C–O, deformation C–H, deformation N–H	32
1400	Symmetric stretching vibration of COO ⁻ group of fatty acids and amino acids	32
1440	Stretching C–H in polysaccharides, pectin	32
1450	Asymmetric CH ₃ bending in proteins	32
1470	CH ₂ bending vibration in lipids and proteins	32
1500	In-plane CH bending vibration from the phenyl ring in phenylalanine, tryptophan or tyrosine	32
1515–1580	Amide II of proteins	14,32,33
1630–1665	β -sheet structure of Amide I of proteins	14,32,33
1683	Unordered random coils and turns of Amide I of proteins	32,33
1689–1698	β -sheet structure of Amide I of proteins	32,33
1700–1708	C=O in thymine	32
1735	C=O in polysaccharides; new COO ⁻ group vibration due to glycosylated human serum albumin	32,34,35
1768–1786	methyl-esterified C=O vibration in IgG COO ⁻ group—glycosylation (IgG with sialylate N-glycans)	33,35–37

Table 2. Assignments for the main vibrational bands of micro-FTIR blood serum (Fig. 2a).

classification data. Histograms in Fig. 6a–c summarize this finding. The positive group presented a broad IgG reactivity index distribution between 1 and 17 while the mix group presented reactivity ranging from 1 to 8. Obviously, the negative group presented a narrow histogram centered on 0.1 value. The outcome from the Z-test ($p < 5\%$) indicated that these IgG distributions come from distinct populations. Interestingly the histograms of signal-to-cutoff from ELISA (Fig. 6e,f,g) showed no evidence that positive and mixing groups belongs to distinct populations ($p > 5\%$ for Z test). Thus, the question that arises concerns the characteristics that gives rise to these distinct populations. Important vibrational bands that contribute to discrimination can help find the answer to this question.

The important vibrational frequencies (VIP) identified by PLS-DA are represented in Fig. 6d,h. For positive and negative discrimination (Fig. 6h and Table 2), β -sheet structure of Amide I of proteins (1689–1698 cm⁻¹) and deoxyribose from DNA bands (840–856 cm⁻¹) contributed to the greatest over-expression in the positive group. Immunoglobulins are heterodimeric proteins composed of two heavy (H) and two light (L) chains. They can be functionally separated into variable (V) domains that bind antigens and constant (C) domains that drives its functions, such as complement activation or binding to Fc receptors³⁸. The basic H2L3 structures consist of the Fc region (fragment, crystallizable) and Fab region (fragment, antigen binding), both composed mainly of β -pleated sheets³⁸. Thus the increased intensity of β -sheet band in the positive group is completely correlated with antibody recruitment in COVID-19 disease. Likewise, the increase in cellular activity due to Sars-Cov-2 inflammatory response explains the observed increase in DNA band intensity. However, for 3 groups (positive/mix/negative) discrimination, a very narrow set of vibrations confined to 1702–1785 cm⁻¹ spectral window is the most important (Fig. 6d). The 1735–1785 cm⁻¹ region had been reported as spectral marker for lipids, C=O cholesteryl esters, and triglycerides³⁵. Besides, the C=O bond for the methyl-esterified carbonyl groups also presents a relatively strong absorption peak at 1730–1760 cm⁻¹³⁶. Also it had been reported that the FTIR spectrum of glycosylated human serum albumin presents a new peak carbonyl group at 1737 cm⁻¹³⁴. Also 1739 cm⁻¹ and 1781 cm⁻¹ bands were reported as predictors for classifying anti-neutrophil cytoplasmic antibodies in sera samples³⁹.

In our case bands in 1702–1785 cm⁻¹ spectral window could be grouped in two classes: (1) those overexpressed in the mix group and subexpressed in the positive group which are both related to IgG glycosylation and (2) those overexpressed in the positive group and underexpressed in the mix group which are both related to thymine. As long as we can argue that the mix group relates to a COVID-19 sub-population with highly glycosylated IgG. As all patients in the present study were mild or oligosymptomatic, the above feature is in agreement with the Chakraborty et al. findings³ which proposed that patients with severe COVID-19 are more likely to have IgG1 with afucosylated Fc glycans as signature. It is well reported that IgG glycosylation can determine whether an antibody glycoform is pro-inflammatory (such as IgG with galactose-deficient N-glycans) or anti-inflammatory (such as IgG with sialylate N-glycans)³⁷. Since 1768–1786 cm⁻¹ region is related to methyl-esterified carbonyl vibration we argue that this band is related to C=O on sialic acid on IgG with sialylate N-glycans. One important point to

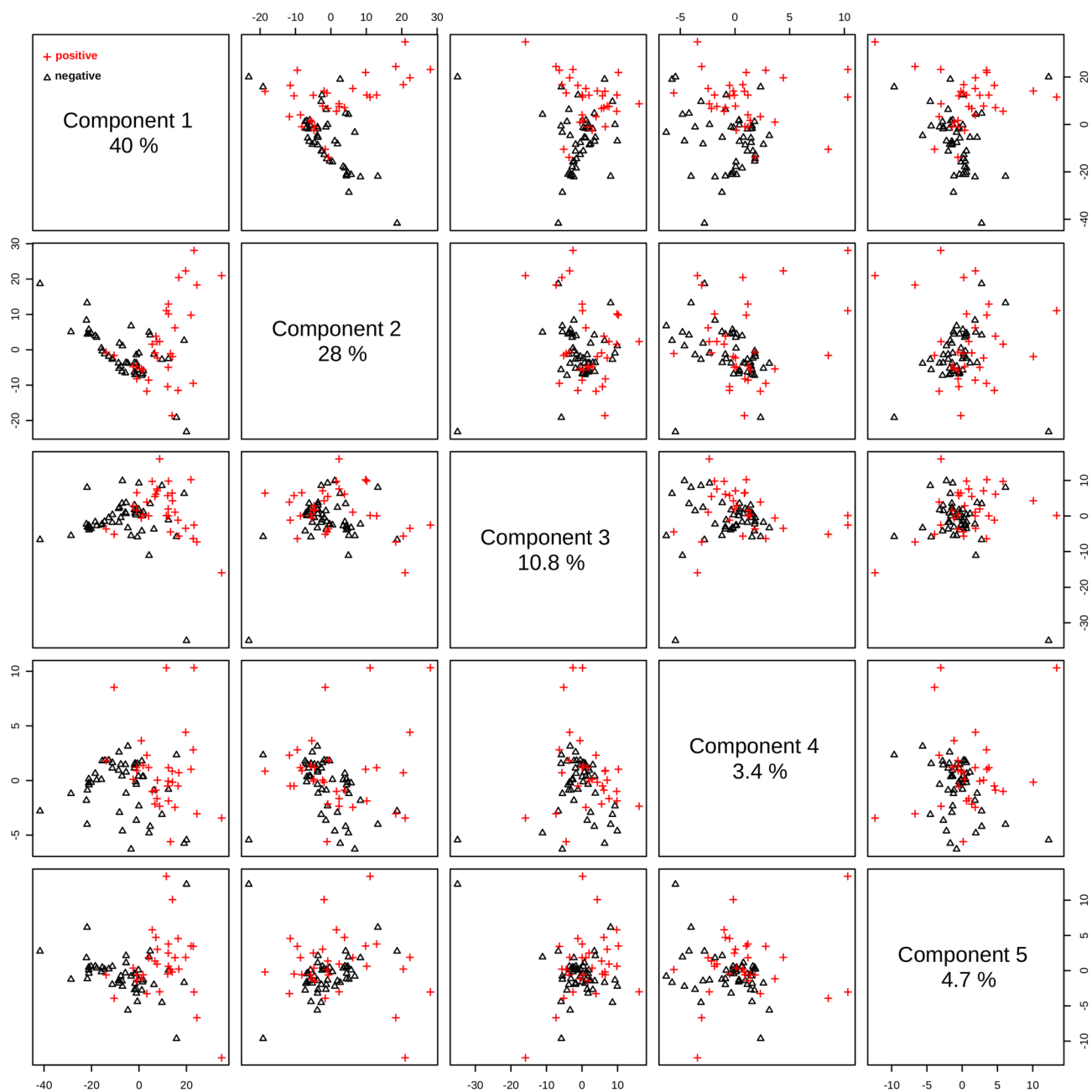


Figure 3. Pairwise score plots for selected PLS-DA components. The explained variance of each component is shown in the corresponding diagonal cell.

clarify is the difference between CLIA and ELISA distribution of Signal-to-cut-off ratio (Fig. 6). It appears that the fluorescence yield of murine monoclonal anti-human IgG anti-body complexed to HRP of CLIA is dependent on the fucosylation of the Fc tail of human IgG. This point needs to be explored in additional experiments.

Concerning the diagnostic performance we mention that spectral windows at 820–890, 1025–1180, and 1685–1727 cm^{-1} presented excellent capability for discriminate negative and positive classes ($\text{AUC} = 0.80\text{--}0.86$, Fig. 7a). The positive and mix classes were also discriminated in excellent grade ($\text{AUC} = 0.80\text{--}0.98$) in spectral windows 850–940, 1015–1170, 1500–1570, 1612–1652, 1695–1735, 1755–1790 cm^{-1} (Fig. 7b). This curve is comparatively noisier respect to the positive/negative case due to the low-N (only 30% of data) in this case. The representative ROC curves for negative and positive (Fig. 7c,d) and mixing and positive (Fig. 7e,f) classes with highest AUC are also shown. The ROC curves for positive/negative (Fig. 7c,d) have binormal curve shapes while mixing/positive ones have straight line shapes due to low-N. The classification boxplots for two discriminating classes presented a significant statistical difference among groups ($p < 0.05$ for t-Student test).

Micro-FTIR reflectance compared to other methods. At this point it is important to compare our results to previous reports concerning blood testing for COVID-19. The usage of ATR-FTIR for analysis of

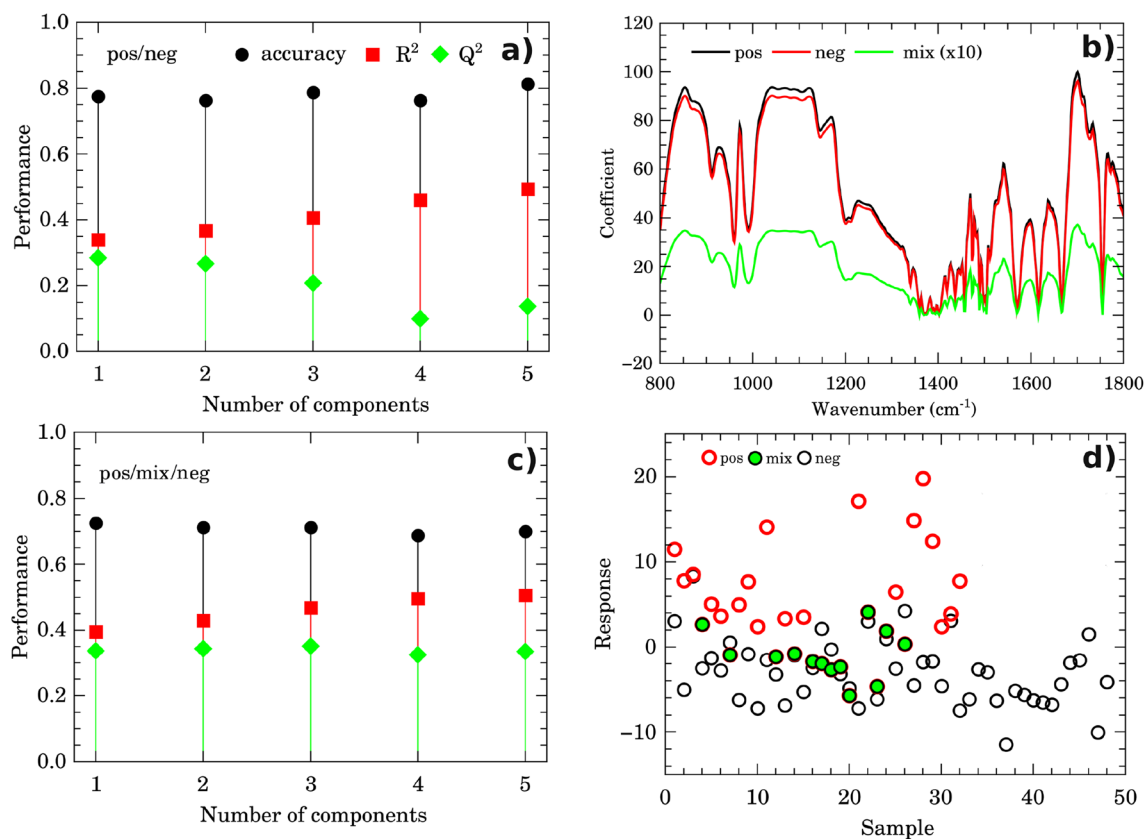


Figure 4. Discrimination performance of micro-FTIR. PLS-DA classification performance using different number of components following accuracy, R^2 and Q^2 criteria for two (positive/negative, (a)) and three (positive/mixture/negative, (c)) groups. Regression coefficients and calculated response in PLS-DA for sample classes are shown in (b,d), respectively.

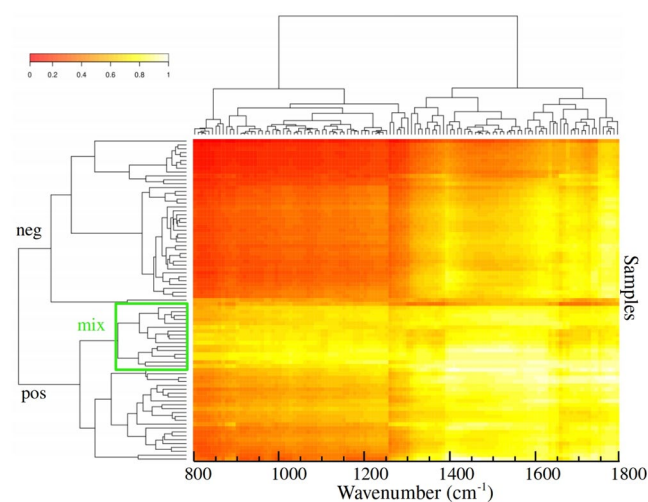


Figure 5. Heatmap for micro-FTIR data. Clustering result shown as a heatmap organized by samples (vertical axis) and wavenumber (horizontal axis). Negative, positive, mix classes grouped into distinct clusters.

plasma from COVID-19 patients have been recently reported in literature. This technique shares a great similarity with our methodology and for this reason a detailed comparison is important.

Banerjee et al.⁴⁰ investigated a cohort of 160 clinicopathologically confirmed SARS-CoV-2 patients using ATR-FTIR. They proposed a plasma processing with 75% ethanol v/v for virus inactivation followed by vortexing and drying over ATR-crystal. The reported AUC was 0.851 when considering the set of spectral and clinical (age, sex, diabetes mellitus, and hypertension) data. The discriminating performance of the spectral data

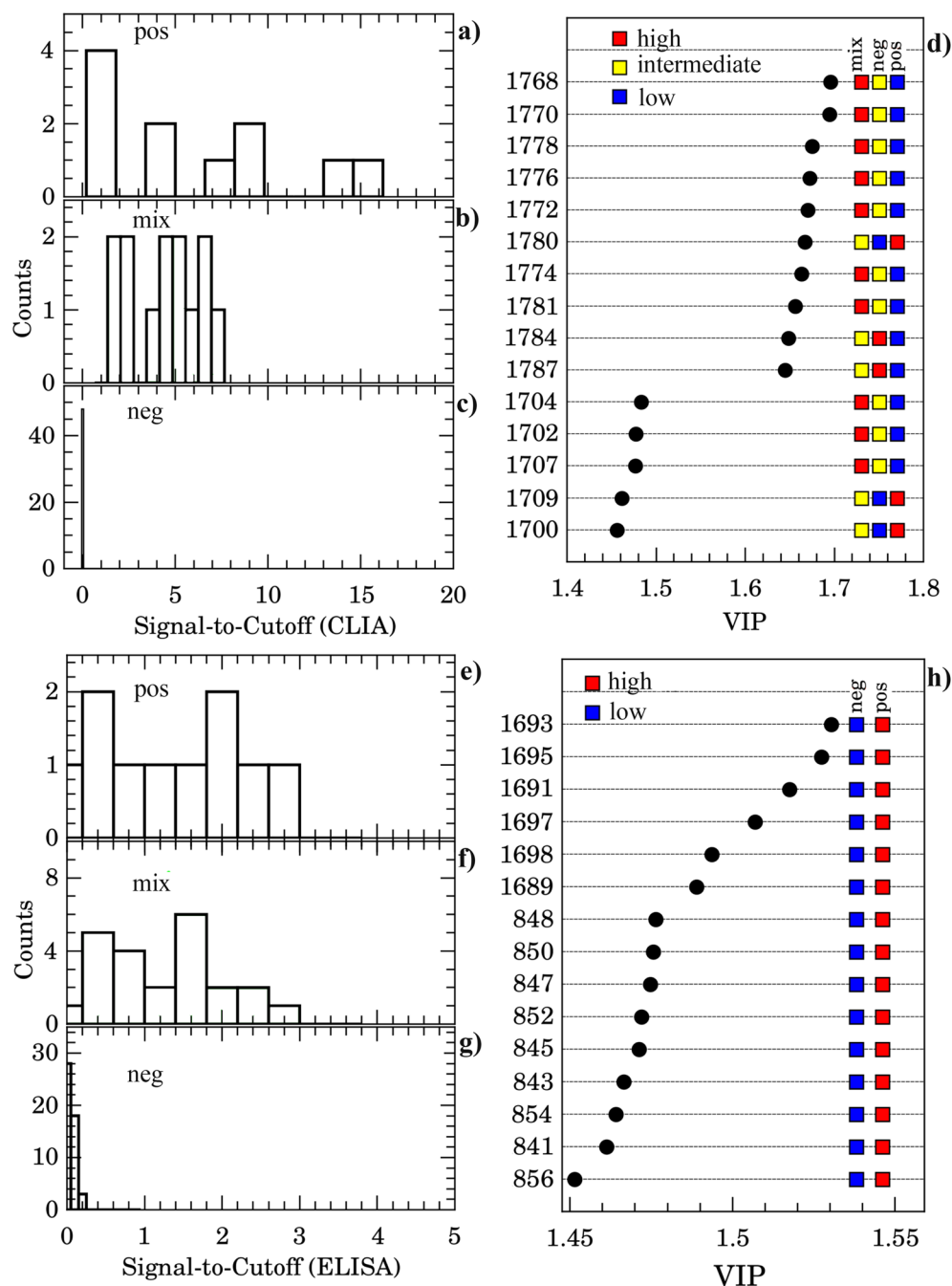


Figure 6. micro-FTIR and CLIA comparison. (a)–(c) Histograms of signal-to-cutoff data of CLIA IgG antibodies against Sars-Cov-2 for positive (a), mix (b), and negative (c) classes as discriminated by micro-FTIR. (d) Important vibrational frequencies (VIP) identified by PLS-DA for three classes classification. (e)–(g) Histograms of signal-to-cutoff data of ELISA IgG antibodies against Sars-Cov-2 for positive (e), mix (f), and negative (g) classes as discriminated by micro-FTIR. (h) VIP for two classes discrimination. Color boxes on the right of (d),(h) indicate the relative intensity (high, intermediate and low) of the corresponding band in each group.

alone was not reported. Zhang et al.⁴¹ investigated by ATR-FTIR blood samples from 20 healthy donors and 76 patients, of which 41 were confirmed with COVID-19, 15 had respiratory viral infections caused by influenza A/B or respiratory syncytial virus (RSV), and 20 were with inflammation-related diseases. In this case the *AUC* for COVID-19 discrimination was reported to be 0.9947. To the best of our knowledge only these two exciting articles reports COVID-19 diagnosis using ATR-FTIR on blood or plasma samples. In spite of diverse sample preparation method and amount of individuals included in the cohort, our *AUC* is very similar to that found by Banerjee et al. and Zhang et al. The key differences between our approach and the ATR-FTIR relies on waiting

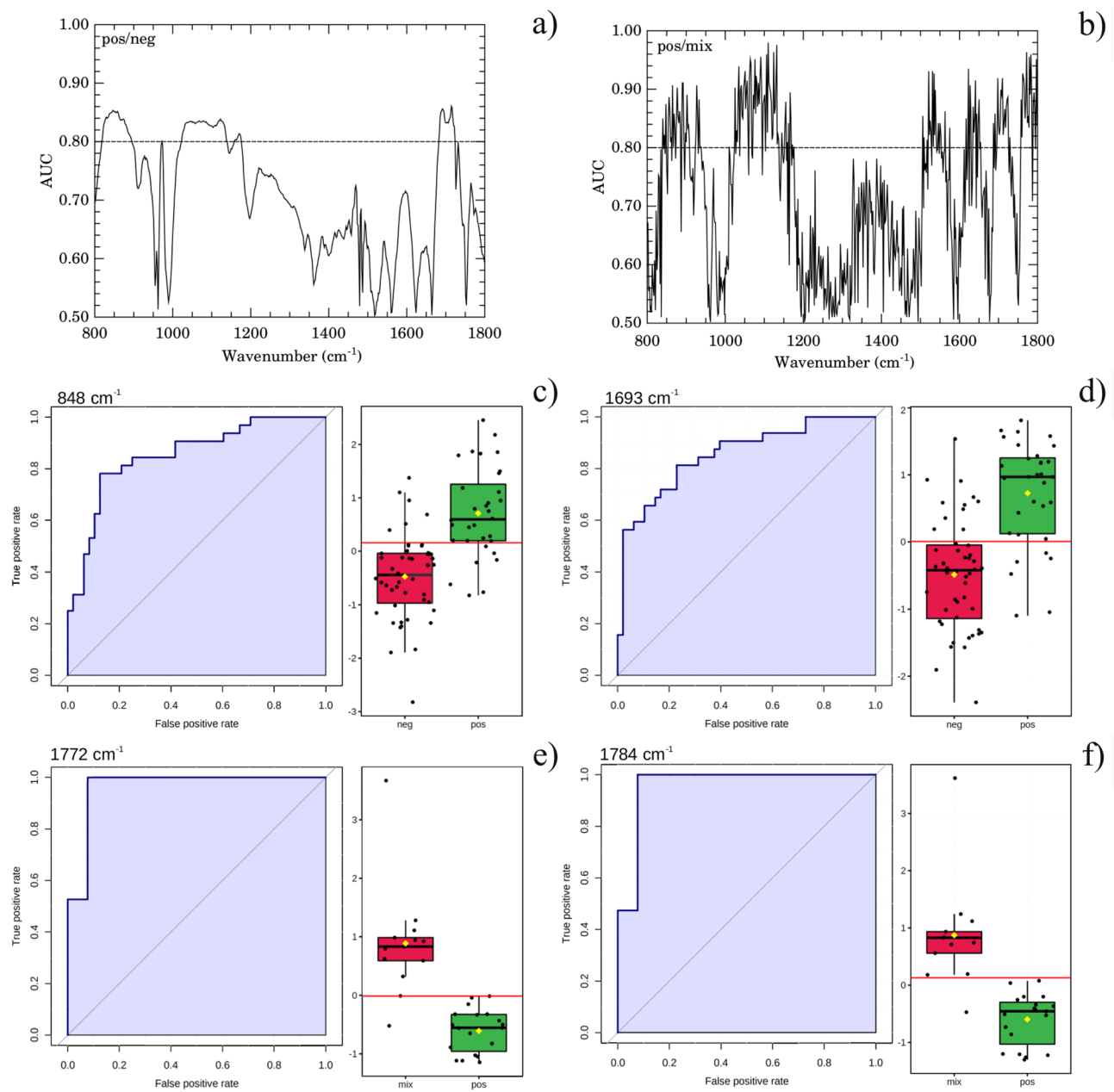


Figure 7. Diagnostic performance of micro-FTIR. Area under receiver operating characteristic (*AUC*) against wavenumber showing those bands with excellent discriminating power ($AUC > 0.80$, dashed horizontal line) for positive/negative (a) and positive/mix (b) classes. Selected representative curves of receiver operating characteristic (ROC) and corresponding classification box-plot of the intensity of the left-handed helix DNA (Z form) (848 cm^{-1} , c), β -sheet structure of Amide I of proteins (1693 cm^{-1} , d), C=O in IgG carbonyl group (1772 and 1784 cm^{-1} in (e,f), respectively) bands. The horizontal red line is the threshold for classification.

time for result outcome, amount of samples which would be tested by row, and COVID-19 fatality capability prediction since we explored in deep the biochemical pieces of information in the FTIR spectra. For spectrum acquisition on ATR-FTIR a single drop of sample need the deposited and dried over the surface of ATR crystal. Then only a single patient could be tested at once. This process demands usually $> 5\text{ h}$ ³⁵. It is not well reported whether the proteins degradation over this time would impact on quality of diagnosis. Moreover the morphological characteristics of the deposited bio-film (heterogeneity, size, wettability, dilution, among others) is dependent of drying conditions as relative moisture and temperature. The quality of spectrum (artifacts, reproducibility) is very sensitive to the morphology of the bio-film. These points impose challenges to the standardization of the method³⁵. On the other hand, our approach enables the preparation of drops of several patients over the substrate. Over an equivalent area of a microplate of 96 wells it is possible the deposition of 500 droplets. Using the microscope re-positioning the IR beam over each patient drop is relatively easy which increases the scale of

Parameter	Micro-FTIR	ATR-FTIR	ELISA	CLIA
Sample preparation	Minimal	Minimal	Complex	Complex
Waiting time for result	< 10 min	> 5 h	2–3 h	2–3 h
Laboratory equipment requirements	Intermediate	Intermediate	High	High
Specialization of human resources for usage	Intermediate	Intermediate	High	High
AUC	> 0.85	> 0.85	> 0.85	> 0.85
Multiplexing capability	Yes	Yes	No	No
COVID-19 fatality prediction	Yes	Depends	No	No
Reagents	Free	Free	Expensive	Expensive
Cost of single test (US\$)	1–10	1–10	50–100	50–100
Cost of main equipment for testing (US\$)	20,000–50,000	20,000–50,000	10,000–30,000	10,000–30,000
Amount of samples tested by row	> 500	1	96	96
Dependence of international logistic and supply chains	Weak	Weak	Strong	Strong
Operator dependence level of reproducibility of outcome	Intermediate	Intermediate	High	High

Table 3. Parameters for maintenance, management, operating, and performance of the main anti-Sars-Cov-2 clinical tests ELISA and CLIA compared to the ATR-FTIR and micro-FTIR.

testing. Banerjee et al. reported the fatality prediction only when considering in the statistical analysis a larger set of data of different kinds as spectra and clinical pieces of information which in some aspect would represent a bias in the statistical test. We were able to discriminate 3 possible groups in an independent way of validation. The clinical data were used to compare and discuss the data but not as input.

Other comparison is with the gold standard method for COVID-19 testing ELISA and also with CLIA method. Both were used to validate our results. Jagtap et al.⁴² evaluated the performance of spike protein antigens for SARS-CoV-2 serology compared four spike proteins: RBD, S1, S2 and a stabilized spike trimer (ST). They used indirect ELISA in serum from COVID-19 patients and pre-2020 samples. ROC curves indicate that ST is the best candidate for serological testing with the highest AUC ($AUC = 0.94$). AUC for IgG was 0.934. Huang et al.⁴³ compared the serological test of SARS-CoV-2 using ELISA test and reported that $AUC(IgM) = 0.812$ while $AUC(IgG + IgM) = 0.983$. The reported AUC for CLIA anti-Sars-CoV-2 test include 0.846⁴⁴, 0.9143⁴⁵, 0.98⁴⁶, and 0.831⁴⁴. Thus we can conclude that the performance of our micro-FTIR approach is in the same level of ELISA and CLIA.

However as summarized by Liu and Rusling⁴⁷, in general the nucleic acid-based tests against COVID-19 have reported high false-negative rates which are dependent of the day of the symptom (38% on the day, 20% 3 days, and 66% 16 days, after of symptom onset, respectively). Factors such as sampling time, viral mutation, inadequate handling, improper storage, transportation of samples, among others had been cited to influence the result. The requirements to perform these tests evolve high workload, skilled personnel for testing and sample collection, special reagent kits, costly centralized infrastructure with specific equipment for process and measured the sample signal, and professional bio safety level (BSL)-2 lab⁴⁷ imposing elevated costs, relatively longer time to delivery of results (2–3 h) and dependence of international chain of suppliers. This last aspect would be disrupted during the pandemic period. One important advantage of the micro-FTIR method is the minimal need of sample preparation which impacts lower cost per testing and relative independence of external suppliers. The time elapsed pipetting, drying, acquiring spectra is typically 10 min. Another aspect to mention is that ELISA and CLIA tests are able to detect a single protein at once. Nevertheless FTIR has multiplexing advantage enabling comparison of several metabolites which increases the quality of prognosis being a suitable tool for precision medicine⁴⁸. Our method enables in principle to probe 500 samples per row while ELISA and CLIA usually are limited to 96 samples. Table 3 summarizes the comparison between all techniques.

Conclusions

Our results showed that micro-FTIR was able to probe key structural aspects of serum IgG from COVID-19 volunteers. The 1702–1785 cm^{-1} spectral window is a spectral marker of the degree of IgG glycosylation, allowing to probe distinctive sub-populations of COVID-19 patients, depending on their degree of severity. Furthermore, the β -sheet structure of Amide I of proteins (1689–1698 cm^{-1}) and deoxyribose from DNA (840–856 cm^{-1}) bands presented most significant contributions to positive and negative discrimination with a specificity of 87.5% and sensibility of 100%. The computed AUC was comparable to the ELISA, CLIA, and other ATR-FTIR methods ($AUC > 0.85$). In summary, overall discrimination of healthy and COVID-19 individuals and severity prediction could also be potentially implemented using micro-FTIR reflectance spectroscopy on blood serum samples due to direct probe of glycosylation degree of IgG. However, experimental efforts need be devoted to isolate immunoglobulins molecules from plasma samples from healthy and COVID-19-ill patients and characterize their spectral properties what would give detailed clues on IgG glycosylation and other structural changes related to pathology.

Received: 17 September 2021; Accepted: 23 February 2022

Published online: 11 March 2022

References

- Mehta, P. *et al.* COVID-19: Consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033–1034. [https://doi.org/10.1016/s0140-6736\(20\)30628-0](https://doi.org/10.1016/s0140-6736(20)30628-0) (2020).
- Hoepfel, W. *et al.* High titers and low fucosylation of early human anti-SARS-CoV-2 IgG promote inflammation by alveolar macrophages. *Sci. Transl. Med.* **13**, eabf8654. <https://doi.org/10.1126/scitranslmed.abf8654> (2021).
- Chakraborty, S. *et al.* Proinflammatory IgG fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73. <https://doi.org/10.1038/s41590-020-00828-7> (2020).
- Luo, H. *et al.* The characterization of disease severity associated IgG subclasses response in COVID-19 patients. *Front. Immunol.* <https://doi.org/10.3389/fimmu.2021.632814> (2021).
- Jermyn, M. *et al.* Intraoperative brain cancer detection with raman spectroscopy in humans. *Sci. Transl. Med.* **7**, 274ra19–274ra19. <https://doi.org/10.1126/scitranslmed.aaa2384> (2015).
- Lin, D. *et al.* Autofluorescence and white light imaging-guided endoscopic raman and diffuse reflectance spectroscopy for in vivo nasopharyngeal cancer detection. *J. Biophoton.* **11**, e201700251. <https://doi.org/10.1002/jbio.201700251> (2018).
- Mitropoulos, K., Katsila, T., Patrinos, G. P. & Pampalakis, G. Multi-omics for biomarker discovery and target validation in biofluids for amyotrophic lateral sclerosis diagnosis. *OMICS J. Integr. Biol.* **22**, 52–64. <https://doi.org/10.1089/omi.2017.0183> (2018).
- Morais, C. L. M., Lima, K. M. G., Singh, M. & Martin, F. L. Tutorial: Multivariate classification for vibrational spectroscopy in biological samples. *Nat. Protocols* **15**, 2143–2162. <https://doi.org/10.1038/s41596-020-0322-8> (2020).
- Barauna, V. G. *et al.* Ultrarapid on-site detection of SARS-CoV-2 infection using simple ATR-FTIR spectroscopy and an analysis algorithm: High sensitivity and specificity. *Anal. Chem.* **93**, 2950–2958. <https://doi.org/10.1021/acs.analchem.0c04608> (2021).
- Nogueira, M. S. *et al.* Rapid diagnosis of covid-19 using ft-ir atr spectroscopy and machine learning. *Sci. Rep.* **11**, 1–13. <https://doi.org/10.1038/s41598-021-93511-2> (2021).
- Kitane, D. L. *et al.* A simple and fast spectroscopy-based technique for covid-19 diagnosis. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-95568-5> (2021).
- Wood, B. R. *et al.* Infrared based saliva screening test for COVID-19. *Angewandte Chemie* **133**, 17239–17244. <https://doi.org/10.1002/ange.202104453> (2021).
- Dogan, A., Gurbanov, R., Severcan, M. & Severcan, F. Coronavac (sinovac) covid-19 vaccine-induced molecular changes in healthy human serum by infrared spectroscopy coupled with chemometrics. *Turk. J. Biol.* **45**, 549. <https://doi.org/10.3906/biy-2105-65> (2021).
- Islam, S. *et al.* Glycation, oxidation and glycoxidation of IgG: A biophysical, biochemical, immunological and hematological study. *J. Biomol. Struct. Dyn.* **36**, 2637–2653. <https://doi.org/10.1080/07391102.2017.1365770> (2017).
- Joyner, M. J. *et al.* Convalescent plasma antibody levels and the risk of death from covid-19. *New England J. Med.* **384**, 1015–1027. <https://doi.org/10.1056/nejmoa2031893> (2021).
- Padoan, A. *et al.* Analytical and clinical performances of five immunoassays for the detection of SARS-CoV-2 antibodies in comparison with neutralization activity. *EBioMedicine* **62**, 103101. <https://doi.org/10.1016/j.ebiom.2020.103101> (2020).
- Shaffaf, T. & Ghafar-Zadeh, E. COVID-19 diagnostic strategies part II: Protein-based technologies. *Bioengineering* **8**, 54. <https://doi.org/10.3390/bioengineering8050054> (2021).
- Cameron, J. M. *et al.* Exploring pre-analytical factors for the optimisation of serum diagnostics: Progressing the clinical utility of ATR-FTIR spectroscopy. *Vib. Spectrosc.* **109**, 103092. <https://doi.org/10.1016/j.vibspec.2020.103092> (2020).
- Varmuza, K. & Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics* (CRC Press, 2016).
- Team, R. C. *et al.* R: A language and environment for statistical computing (2013).
- Lieber, C. A. & Mahadevan-Jansen, A. Automated method for subtraction of fluorescence from biological raman spectra. *Appl. Spectrosc.* **57**, 1363–1367. <https://doi.org/10.1366/000370203322554518> (2003).
- Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1h NMR metabolomics. *Anal. Chem.* **78**, 4281–4290. <https://doi.org/10.1021/ac051632c> (2006).
- Wehrens, R. & Mevik, B.-H. pls: Partial least squares regression (pls) and principal component regression (pcr), r package version 2.1-0. <https://doi.org/10.18637/JSS.V018.I02> (2007).
- Bijlsma, S. *et al.* Large-scale human metabolomics studies: A strategy for data (pre-) processing and validation. *Anal. Chem.* **78**, 567–574. <https://doi.org/10.1021/ac051495j> (2005).
- Szymańska, E., Saccenti, E., Smilde, A. K. & Westerhuis, J. A. Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **8**, 3–16. <https://doi.org/10.1007/s11306-011-0330-3> (2011).
- Zou, K. H., O'Malley, A. J. & Mauri, L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* **115**, 654–657. <https://doi.org/10.1161/circulationaha.105.594929> (2007).
- Ma, H., Bandos, A. I. & Gur, D. On the use of partial area under the ROC curve for comparison of two diagnostic tests. *Biomet. J.* **57**, 304–320. <https://doi.org/10.1002/bimj.201400023> (2014).
- Justel, A., Peña, D. & Zamar, R. A multivariate kolmogorov-smirnov test of goodness of fit. *Stat. Probab. Lett.* **35**, 251–259. [https://doi.org/10.1016/s0167-7152\(97\)00020-5](https://doi.org/10.1016/s0167-7152(97)00020-5) (1997).
- Markowski, C. A. & Markowski, E. P. Conditions for the effectiveness of a preliminary test of variance. *Am. Stat.* **44**, 322. <https://doi.org/10.2307/2684360> (1990).
- Sanyaolu, A. *et al.* Comorbidity and its impact on patients with COVID-19. *SN Compr. Clin. Med.* **2**, 1069–1076. <https://doi.org/10.1007/s42399-020-00363-4> (2020).
- Krebs, H. A. Chemical composition of blood plasma and serum. *Ann. Rev. Biochem.* **19**, 409–430. <https://doi.org/10.1146/annurev.bi.19.070150.002205> (1950).
- Movasaghi, Z., ur Rehman, S. & Rehman, D. I. Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **43**, 134–179. <https://doi.org/10.1080/05704920701829043> (2008).
- Ahmad, S., Moinuddin, K. R. H. & Ali, A. Physicochemical studies on glycation-induced structural changes in human IgG. *IUBMB Life* **64**, 151–156. <https://doi.org/10.1002/iub.582> (2012).
- Raghav, A., Ahmad, J. & Alam, K. Nonenzymatic glycosylation of human serum albumin and its effect on antibodies profile in patients with diabetes mellitus. *PLoS One* **12**, e0176970. <https://doi.org/10.1371/journal.pone.0176970> (2017).
- Cameron, J. M. *et al.* Developing infrared spectroscopic detection for stratifying brain tumour patients: Glioblastoma multiforme vs. lymphoma. *Analyst* **144**, 6736–6750. <https://doi.org/10.1039/c9an01731c> (2019).
- Zhu, R. *et al.* Pectin oligosaccharides from hawthorn (*Crataegus pinnatifida* bunge. var. major): Molecular characterization and potential antiglycation activities. *Food Chem.* **286**, 129–135. <https://doi.org/10.1016/j.foodchem.2019.01.215> (2019).
- Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462. <https://doi.org/10.1038/nrm3383> (2012).
- Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52. <https://doi.org/10.1016/j.jaci.2009.09.046> (2010).
- Durlík-Popińska, K. *et al.* Correlations between autoantibodies and the atr-ftir spectra of sera from rheumatoid arthritis patients. *Sci. Rep.* **11**, 1–11. <https://doi.org/10.1038/s41598-021-96848-w> (2021).

40. Banerjee, A. *et al.* Rapid classification of COVID-19 severity by ATR-FTIR spectroscopy of plasma samples. *Anal. Chem.* **93**, 10391–10396. <https://doi.org/10.1021/acs.analchem.1c00596> (2021).
41. Zhang, L. *et al.* Fast screening and primary diagnosis of COVID-19 by ATR-FT-IR. *Anal. Chem.* **93**, 2191–2199. <https://doi.org/10.1021/acs.analchem.0c04049> (2021).
42. Jagtap, S. *et al.* Evaluation of spike protein antigens for SARS-CoV-2 serology. *J. Virol. Methods* **296**, 114222. <https://doi.org/10.1016/j.jviromet.2021.114222> (2021).
43. Huang, S. *et al.* Rapid detection of COVID-19 by serological methods and the evaluation of diagnostic efficacy of IgM and IgG. *Clin. Lab.* <https://doi.org/10.7754/clin.lab.2020.200617> (2020).
44. Soleimani, R. *et al.* Clinical usefulness of fully automated chemiluminescent immunoassay for quantitative antibody measurements in COVID-19 patients. *J. Med. Virol.* **93**, 1465–1477. <https://doi.org/10.1002/jmv.26430> (2020).
45. Mendoza, R. *et al.* Correlation of automated chemiluminescent method with enzyme-linked immunosorbent assay (ELISA) antibody titers in convalescent COVID-19 plasma samples: Development of rapid, cost-effective semi-quantitative diagnostic methods. *J. Blood Med.* **12**, 157–164. <https://doi.org/10.2147/jbm.s296730> (2021).
46. Fabián, C. L. R. & Briceño, L. U. Anti-SARS-cov-2 IgA in current scenario of IgM and IgG rapid test: A new alternative for the diagnostic of COVID-19. *SN Compr. Clin. Med.* **2**, 2167–2169. <https://doi.org/10.1007/s42399-020-00551-2> (2020).
47. Liu, G. & Rusling, J. F. COVID-19 antibody tests and their limitations. *ACS Sens.* **6**, 593–612. <https://doi.org/10.1021/acssensors.0c02621> (2021).
48. Zhou, A. *et al.* Is precision medicine relevant in the age of COVID-19?. *Genet. Med.* **23**, 999–1000. <https://doi.org/10.1038/s41436-020-01088-4> (2021).

Acknowledgements

The authors would like to thank the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – 311146/2015-5 and 307718/2019-0) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Projeto CAPES-Pandemias - 88881.504639/2020-01) for the financial support. The authors would also thank the experimental resources provided by Multiuser Central Facilities at UFABC (CEM/UFABC) and the financial support by UFABC (Projeto Enfrentamento à COVID-19/reitoria).

Author contributions

C.C.S.B. performed the spectroscopy experiments; K.C.R.M and M.B.R. conducted the samples collecting and analyzed the clinical data; J.F.G., A.P.M.A.S., V.L.T., V.A.L., N.K.J., and R.M.Z. performed the biomolecular tests; J.D.A., P.H.B.-S., and J.A.L. discussed the biomolecular results; H.S.M. wrote the main manuscript texts and prepared the figures. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022