Scientific Reports

https://doi.org/10.1038/s41598-025-34555-6

Article in Press

Assessment of the effectiveness of serum-infrared spectroscopy in conjunction with multivariate analysis methods for atherosclerosis diagnosis

Received: 21 May 2025

Accepted: 29 December 2025

Published online: 03 January 2026

Cite this article as: Cevik D., Telkoparan-Akillilar P. & Yonar D. Assessment of the effectiveness of serum-infrared spectroscopy in conjunction with multivariate analysis methods for atherosclerosis diagnosis. *Sci Rep* (2026). https://doi.org/10.1038/s41598-025-34555-6

Dilek Cevik, Pelin Telkoparan-Akillilar & Dilek Yonar

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

ARTICLE IN PRESS

Assessment of the effectiveness of serum-infrared spectroscopy in conjunction with multivariate analysis methods for atherosclerosis diagnosis

Dilek Çevik[1], Pelin Telkoparan-Akillilar[2], Dilek Yonar[3,4*]

[1]Department of Molecular Biology and Genetics, Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Türkiye

[2]Department of Medical Biology, Faculty of Medicine, Gazi University, Ankara, Türkiye

[3]Department of Biophysics, Faculty of Medicine, Yuksek Ihtisas University, Ankara, Türkiye

[4]Department of Physics, Faculty of Engineering, Atılım University, Ankara, Türkiye

ARTICLE IN PRESS

***Corresponding Author:** Dilek Yonar

Atılım University

Faculty of Engineering

Department of Physics

06836, Ankara, Türkiye

e-mail: dilekyonar81@gmail.com / dilek.yonar@atilim.edu.tr

ARTICLE IN PRESS

Abstract

Atherosclerosis is a progressive disease characterized by the lipid accumulation and fibrous elements in large and medium-sized arteries, and remains a leading cause of death worldwide. A deeper understanding of its molecular nature is critical for developing novel strategies for the prevention, diagnosis, and treatment. This study evaluates attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectroscopy combined with multivariate analysis techniques to differentiate and classify atherosclerosis by identifying disease-specific spectral variations. Spectral analyses indicated statistically significant differences in lipid ($p<0.0001$), protein ($p<0.01$), nucleic acid ($p<0.0001$), and glucose ($p<0.0001$) content of serum samples in the atherosclerosis group compared to controls. Atherosclerosis patients exhibit altered lipid metabolism, marked by a decrease in saturated lipids and an increase in unsaturated lipids compared to healthy individuals. Additionally, elevated levels of protein, RNA, glucose, and conformational changes in DNA were key spectral features, distinguishing atherosclerosis from controls. Principal component analysis (PCA) successfully differentiated patients from controls, while classification models based on linear discriminant analysis (LDA) and support vector machine (SVM) achived accuracies of 96.61% and 93.22%, respectively. The ability of FTIR spectroscopy to detect subtle biochemical alterations suggests its potential for early diagnosis. These molecular markers might appear before clinical symptoms, highlighting the method's potential for future screening, pending validation in at-risk or preclinical cohorts.

Keywords: atherosclerosis; blood serum; spectral biomarkers; ATR-FTIR spectroscopy; multivariate analysis.

ARTICLE IN PRESS

ARTICLE IN PRESS

Introduction

Atherosclerosis is a leading cause of cardiovascular morbidity and mortality worldwide, imposing a significant burden on healthcare systems and patients' quality of life. This chronic inflammatory disease is characterized by the impairment of normal vessel function due to the accumulation of plaque within the arterial walls.[1,2] Lipids, cellular debris, immune cells, calcium, and extracellular matrix components accumulate over time, resulting in arterial narrowing, reduced blood flow, and increased risk of life-threatening events such as thrombosis, stroke, and myocardial infarction.[2,3] Atherosclerosis often progresses silently, with no symptoms until severe complications arise, making it one of the primary causes of premature death worldwide.[3] Even if the patients survive atherosclerosis, they often suffer from long-term health problems that significantly reduce their quality of life, placing a substantial economic burden on health care systems. Various imaging techniques, such as ultrasound, CT angiography, and MRI, are currently employed to detect and monitor the progression of atherosclerosis.[4] However, these imaging techniques are frequently expensive, necessitate specialized equipment and experienced professionals, and may exhibit insufficient sensitivity to detect early-stage biochemical alterations that precede visible plaque formation. Furthermore, they may not always be suitable for large-scale screening or regular monitoring. According to The American Heart Association (AHA), detecting subclinical atherosclerosis would open the way for prophylactic measures to prevent irreversible vascular damage.[5] Early detection of atherosclerosis is crucial for initiating effective preventive and therapeutic interventions. Therefore, it is essential to develop alternative diagnostic tools that provide accurate, non-invasive or minimally-invasive, rapid, cost-effective, operator-independent and sensitive to early molecular changes.

Fourier transform infrared spectroscopy (FTIR) is a promising technique for analyzing various biological materials, including tissue sections, cytological and histological samples or biological fluids.[6] Attenuated total reflectance (ATR) mode enables a rapid and easy way to obtain high quality IR spectra for qualitative analysis with no sample preparation and also provides highly reproducible results for a wide range of sample types.[7] ATR-FTIR spectroscopy offers a reagent-free and real-time biochemical profiling of biological samples such as serum and plasma, which are rich in disease-relevant biomarkers and easily accessible for the early detection of various diseases.[6–9] It has already been applied in the diagnosis of cancer, neurodegenerative diseases and metabolic diseases.[7,9–16] Additionally, IR spectral measurements combined with multivariate analysis and machine learning methods can be

ARTICLE IN PRESS

successfully used to identify spectral biomarkers and classify diseases by extracting substantial information from complex spectral data.[7,10,17,18]

There is a great deal of literature on vibrational spectroscopic evaluation of human atherosclerotic plaques.[19–23] Li et al. (2003) conducted a study showing the evolution of lipid bands, fibrous tissue bands and phosphate calcification bands within the plaques, as observed in spectra taken from various regions of the plaques.[21] Anastassopoulou et al. (2022) proposed that FTIR spectra reveal key molecular changes in atherosclerosis, including aldehyde formation resulting from lipid and protein peroxidation, the accumulation of advanced glycation end products (AGEs), and calcium salt deposition, all of which contribute to carotid artery stenosis and plaque formation.[22] Marzec et al. (2015) applied vibrational spectroscopy in the biochemical analysis and imaging of the key plaque features, including vascular wall components, markers of endothelial dysfunction, and heart valve stenosis.[24] More recently, the focus has shifted toward serum-based analysis for early-stage detection, because serum samples of atherosclerosis patients will display unique molecular and biochemical alterations when compared to healthy individuals. Infrared spectroscopy reflects these alterations in the absorption patterns associated with the modified lipid, protein, carbohydrate, and nucleic acid compositions in the samples. Peters et al. (2017) conducted an exploratory pilot study to investigate the use of serum infrared spectroscopy combined with neural network analysis for detecting and differentiating secondary complications of atherosclerosis. They suggested this combination as a promising diagnosing method, noting that the multifactorial development of atherosclerosis leads to widespread changes in serum composition, which can be detected with high sensitivity and specificity using FTIR analysis supported by neural networks.[25] Kęsik et al. (2024) further integrated ATR-FTIR spectroscopy with machine learning to investigate serum-based markers indicative of atherosclerotic carotid stenosis (ACS). Their results demonstrated substantial spectral differences between healthy individuals and ACS patients, particularly involving phospholipid and lipid-associated vibrational modes.[26]

A significant number of researches primarily focus on the characterization of tissue samples (e.g., carotid plaques) by FTIR spectroscopy. The present study proposes a minimally-invasive and clinically-translatable diagnostic framework utilizing ATR-FTIR spectroscopy combined with advanced multivariate analysis and machine learning methods for the diagnosis of atherosclerosis via serum sample. This work differs from previous serum-based studies by the difference in the disease type and introducing the different machine learning algorithms. Our approach not only differentiates atherosclerosis patients from healthy controls with high accuracy, but also enables the biochemical interpretation of key spectral features associated

ARTICLE IN PRESS

with disease etiology. Our approach contributes to the development of clinically applicable, operator-independent, and cost-effective screening tools for cardiovascular diseases by addressing both the analytical and interpretative aspects of FTIR-based diagnosis.

Results and Discussion

This study was performed to elucidate the molecular basis of atherosclerosis concerning biomolecular content and structure. This method is expected to facilitate the identification of spectral biomarkers. Furthermore, ATR-FTIR spectroscopy combined with multivariate analysis methods was employed to develop an innovative, rapid, and expedient method for diagnosing atherosclerosis and differentiating it from healthy individuals using blood serum.

Fig. 1 presents the FTIR spectra of serum samples from both controls and patients, focusing on the 3015–2800 and 1800–800 cm^{-1} regions (The full-range raw spectra are provided as Fig. S1 in the supplementary material). These regions display characteristic absorption bands corresponding to the functional groups of biomolecules, including lipids, proteins, carbohydrates and nucleic acids (RNA, DNA). The assignments of the absorption bands in serum samples were summarized in our previous manuscript.[7] The difference spectra shown in the subpanels of Fig. 1 provide a clearer representation of the disparities between the atherosclerosis and control groups. Significant spectral differences were observed in the serum biomolecules from individuals with atherosclerosis compared to the controls (Fig. 1). Specifically, the pronounced differences in the 1300–800 cm^{-1} range, driven by changes in nucleic acids, proteins, and carbohydrates, indicate atherosclerosis-associated metabolic alterations (Fig. 1, Fig. S1 and Fig. S2). A comprehensive spectral analysis is essential to accurately identify the alterations induced by atherosclerosis. A more detailed information about spectral band analyses can be found in our previous study.[7] As discussed in our prior study, band area ratios were employed instead of raw area values to eliminate experimental distortions and ensure accurate concentration measurements of biomolecules.

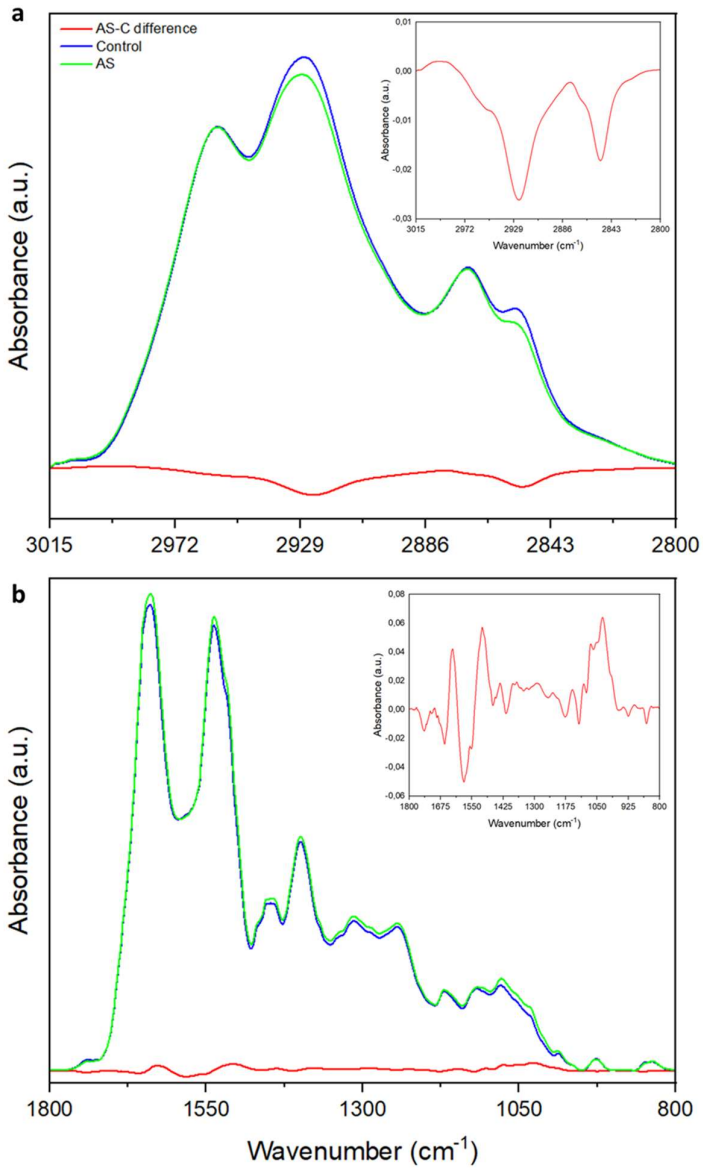


Figure 1. Average serum spectra of the control and AS groups in a) the 3015–2800 and b) the 1800-800 cm^{-1} regions. The spectra were normalized to the amide A band located at 3286 cm^{-1}. Difference IR spectra for both regions, shown in the subpanels, were obtained by the subtraction of the control spectrum from AS patients' spectra. Control (blue), AS (green) and difference (red) spectra are displayed.

IR absorptions in the 3020–2800 cm^{-1} region, which include olefinic =CH stretching and the antisymmetric and symmetric stretching modes of methylene ($–CH_2$) and methyl ($–CH_3$) groups are sensitive indicators of lipid structure and composition. These spectral features were analyzed to assess changes in lipid saturation and chain structure between control and atherosclerosis (AS) groups. Specifically, the antisymmetric stretching at (2922 cm^{-1}) and

ARTICLE IN PRESS

symmetric stretching (2852 cm^{-1}) vibrations of methylene groups, originating from the long hydrocarbon chains of saturated lipids, were employed to quantify saturated lipid content, as indicated by the total integrated area under these peaks. A significant reduction in saturated lipid content was observed in the AS group compared to controls (6.55 vs. 6.74, $p < 0.01$). The unsaturation index, defined by the ratio of unsaturated to saturated lipids, was significantly higher in the serum samples of AS patients compared to controls (Fig. 2a). These findings indicate a compositional shift toward increased lipid unsaturation in atherosclerosis, which may influence membrane fluidity, susceptibility to lipid peroxidation, and overall lipid metabolism. This shift aligns with previous findings of lipid remodeling in cardiovascular disease. A cross-sectional study has demonstrated that serum lipid profiles in individuals with ischemic heart disease typically show elevated total cholesterol, LDL-C, and triglycerides, reduced HDL-C levels, reflecting systemic lipid dysregulation.[27] More detailed lipidomic analyses have revealed that specific plasma lipid species are markedly altered in coronary artery disease and serve as predictors of long-term clinical outcomes.[28] Similarly, alterations in the lipid composition of circulating LDL have been documented in atherosclerosis, implying that lipoprotein functionality is affected by lipid saturation and oxidation states.[29] These findings support the idea that the increase in unsaturation observed in our study reflects a broader pattern of lipid remodeling associated with atherosclerotic disease progression, which may influence lipoprotein behavior, membrane properties, and inflammatory responses.

Additionally, the lower saturated lipid content in the blood may be attributed to differences in lipid degradation pathways between patients and controls.[30] Supporting this, the methyl/methylene area ratio was higher in atherosclerosis group compared to controls, indicating the presence of relatively shorter chain and/or more branched lipids. Such structural characteristics may suggest increased lipid degradation (Fig. 2b). Shorter-chain fatty acids are known to trigger immune responses and promote chronic inflammation, which are key contributors to atherosclerosis progression.[31] As a result, elevated amounts of shorter-chain fatty acids in patient samples may serve as a marker as well as a mediator of disease. Further insight into lipid structural disorder was obtained by determining the shifts in the wavenumbers of the antisymmetric and symmetric CH_2 stretching vibrations, which reflects the order/disorder state of lipid acyl chains.[7,15] In AS samples, the peak positions of the CH_2 antisymmetric band shifted significantly to higher wavenumber (2930.11 cm^{-1} in AS vs. 2929.19 cm^{-1} in controls, $p < 0.01$), indicating increased conformational disorder in the lipid hydrocarbon chains. These modifications confirm the presence of altered lipid packing and flexibility, which are characteristic of a more disordered membrane environment in atherosclerosis.

ARTICLE IN PRESS

Finally, the absorption at 1739 cm^{-1}, corresponding to the stretching carbonyl vibrations C=O of ester carbonyl groups, indicates the triglycerides and cholesterol ester content in the sample and is monitored to assess lipid peroxidation by free radical activity.[32] However, no significant change was observed in this band, indicating that hyperoxidation was not significantly elevated in our cohort. This is consistent with clinical data from the same group of patients, which showed no significant differences in mean LDL and triglyceride levels compared to controls (Table 1).[33]

Fig. 2c and 2d illustrate the total protein amount and structural and conformational alterations of the proteins, respectively. Variations in total protein content are determined by measuring the total areas under the amide I (1639 cm^{-1}) and amide II (1538 cm^{-1}) bands. A significant increase ($p < 0.01$) in total protein content associated with atherosclerosis was observed (Fig. 2c), highlighting alterations in protein metabolism in atherosclerosis. The amide I/amide II area ratio, along with shifts in the frequencies of these amide bands, demonstrates the structural and conformational changes of the proteins.[7,15,34] For instance, a slight non-significant decrease in the wavenumber of the amide I band in atherosclerosis suggests alterations in protein conformation (the wavenumbers of the amide I band were 1639.73 cm^{-1} for controls and 1638.91 cm^{-1} for AS, p=0.13). Additionally, the observed increase in the amide I/amide II area ratio represents structural variations in proteins (Fig. 2d). This result is consistent with changes in protein structure associated with disease states. These results align with those presented in the study of Kotoulas et al. (2017), who employed FTIR spectroscopy to examine coronary arteries in diabetic individuals with atherosclerosis. Their research demonstrated similar shifts in the amide I and II bands toward lower frequencies, signifying a conformational transition from α-helical to β-sheet structures.[35] The presence of β-sheet-rich structures is particularly noteworthy as it indicates protein misfolding. This misfolding is frequently induced by oxidative stress and chronic inflammation, both of which play key roles in the pathogenesis of atherosclerosis and diabetes.[36]

The bands related to phosphate antisymmetric and symmetric stretching of PO_2^- groups in nucleic acids and phospholipids are located at 1243 and 1080 cm^{-1}, respectively. Protein phosphorylation was quantified using the area ratio of the band located at 1243 cm^{-1} to the band at 2958 cm^{-1} , as previously described.[34] A significant decrease in this parameter was observed in the AS group ($p < 0.05$; Fig. 2e), suggesting dysregulation of protein phosphorylation during atheroma formation. These findings align with the understanding that posttranslational modifications, particularly protein phosphorylation, are dysregulated during atherosclerosis progression.[37,38] Novel techniques have been developed to detect changes in protein

ARTICLE IN PRESS

phosphorylation in atherosclerosis for diagnostic purposes.[39] Accordingly, the detection of phosphorylation levels using FTIR could serve as an alternative method to approach to assess protein phosphorylation levels.

Specific bands corresponding to DNA and RNA are observed around 1121, 991, 926, 852 and 835 cm^{-1}. Fig. 2f shows the RNA concentration calculated as the area ratio of the RNA related band at 1121cm^{-1} to the total nucleic acid, which demonstrated a significant increase ($p < 0.0001$). This result highlights an increase in RNA concentration in atherosclerosis, consistent with numerous studies that have investigated specific RNAs and microRNAs. These studies report differential levels of circulating RNAs or microRNAs in an effort to establish a functional link between the abundance or scarcity of these RNAs in the blood samples of atherosclerosis patients.[40–48]

Furthermore, different conformations of DNA helix were detected in AS samples, as evidenced by an increase in the B form of DNA in the patients. In contrast, no significant difference was observed in the Z form of DNA between AS and control samples (Fig. 2g). DNA methylation is known to induce a shift from the B form to the Z form of DNA,[49] and abnormal changes in the DNA methylation have been reported in various diseases including atherosclerosis.[50] These findings are consistent with previous reports suggesting that conformational changes in DNA may serve as an indicator to distinguish atherosclerosis samples from controls. Fig. 2h also shows the change in the A-form conformation of DNA, determined by area ratio of the DNA related band at 852 cm^{-1} to the total nucleic acid, with a significant decrease observed in this area ratio. Since atherosclerosis is characterized by oxidative stress and inflammation, DNA damage, including strand breaks, base modifications, or damaged DNA conformations may occur.[51,52] These modifications may cause the disruption of the typical A-form conformation. Quantitation of glucose is often achieved using the spectral band at 1030 cm^{-1}, which corresponds to the C-O stretching of carbohydrates, particularly glucose.[7,53] Fig. 2i demonstrates the glucose level with respect to the protein amount. This band area ratio is obtained by taking the ratio of absorbance areas under the 1030 cm^{-1} band and the total area under the amide I and II bands. This ratio was higher AS group as compared to control. Extensive research has focused on the relationship between elevated blood glucose levels (hyperglycemia) and atherosclerosis, demonstrating a direct correlation between glucose levels and the progression of atherosclerosis.[54–58] Elevated glucose levels stimulate the formation of reactive oxygen species (ROS), which in turn triggers inflammatory responses.[59] The increased area of peaks related to RNA and carbohydrate vibrations suggests enhanced RNA synthesis

and glucose metabolism in the diseased group, consistent with inflammatory and oxidative stress processes associated with atherosclerosis.

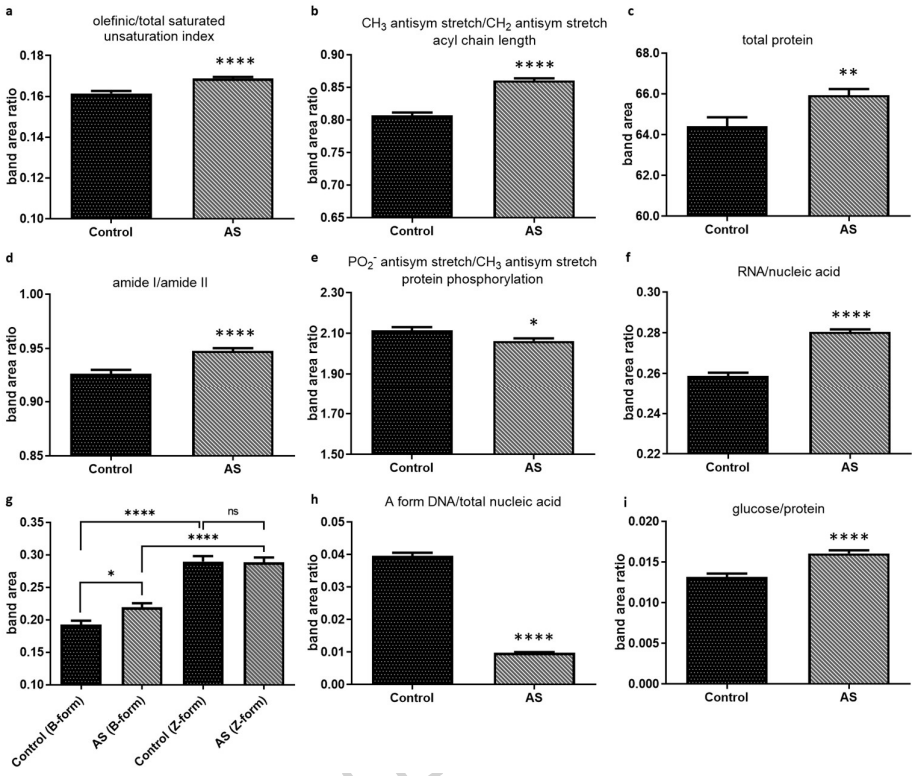


Figure 2. Lipid, Protein, and Nucleic Acid Alterations in AS vs. Control Groups. Bar graphs showing: a) olefinic/total saturated lipid area ratio, b) methyl/methylene area ratio, c) total protein area, d) amide I/ amide II area ratio, e) PO_2^- antisymmetric stretching/CH_3 antisymmetric stretching area ratio, f) RNA/nucleic acid area ratio, g) comparison of the bands areas B-form and Z-form DNA, h) A-form DNA/nucleic acid ratio, and i) glucose/protein area ratio of control and AS groups (The degree of significance for the comparison of the AS group with respect to the control group was denoted as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

Supervised multivariate analysis methods such as discriminant analysis and support vector machines (SVM) have gained significant attention in disease diagnosis. These advanced techniques utilize data-driven algorithms to extract meaningful patterns from complex datasets, effectively classifying diverse disease states. Moreover, they enable the identification of specific biomarkers and disease-related patterns that are often difficult to detect using traditional diagnostic methods. In this research, we aimed to develop a fast and reliable diagnostic approach that encompasses a range of spectral variables potentially related to atherosclerosis using FTIR spectral data of blood serum samples.

ARTICLE IN PRESS

Initially, we utilized PCA, an unsupervised method, to explore the variances within the dataset and to visualize potential trends between the control and AS groups. PCA was used to identify patterns and spectral regions that contribute most to group differences. The optimal spectral region was identified as 1300-800 cm^{-1} from the loading plots. The results for this region are presented in the score and loading plots shown in Fig. 3. The score plots in Fig. 3a, displayed in the PC1-PC2 plane, reveal a trend toward separation between the AS and the control group serum samples, with the main variation observed along PC1, which accounts for 43% of the total variance (Fig. 3a). The corresponding loading plots (Fig. 3b) emphasize specific molecular vibrations contributing to this trend, including C-O stretching of ribose (1121 cm^{-1}), PO_2^- symmetric stretching of nucleic acids (1090, 1079 cm^{-1}), C-O stretching of carbohydrate residues (glucose) (1030 cm^{-1}), RNA backbone vibration (991 cm^{-1}) and the functional groups related to DNA conformations (926, 852 cm^{-1}). Along PC1, the discriminating positively correlated loadings are found for the control, whereas negatively correlated loadings are characteristic for AS (Fig. 3b, Fig. S2). The loading plot shown in Fig. 3b indicates that the largest changes in the composition of serum samples due to inducing AS (positive scores) are associated with nucleic acid and carbohydrate components (negative loading at about 1079 cm^{-1} along with others at 1030 and 991 cm^{-1}) with respect to controls (negative scores). The PO_2^- group vibrations at around 1079 cm^{-1} and C-O vibrations at 1030 cm^{-1} serve as the principal biomarkers for distinguishing AS patients from controls along the PC1 axis. The largest changes in the composition of control serum samples (negative scores) are associated with nucleic acid components (positive loading at about 1090 cm^{-1} along with others at 1121, 1000, 926 and 852 cm^{-1}). The strong positively correlated band at 1090 cm^{-1} and thus specific for controls is also attributed to the PO_2^- symmetric stretching, and its shift and opposite correlation compared to the 1079 cm^{-1} band is a strong indicator of subtle structural differences in the phosphate group between two groups.

Consequently, PCA score plots show a general trend toward separation between AS and control samples, with relevant molecular vibrations emerging as potential biomarkers. The existence of several peaks in the loading plots of the discriminating principal components and the compatibility of their spectral positions with those of the difference plots, validated the discriminative power of infrared spectra. These findings provide a basis for developing non-invasive diagnostic methods and further understanding the biochemical mechanisms of atherosclerosis.

ARTICLE IN PRESS

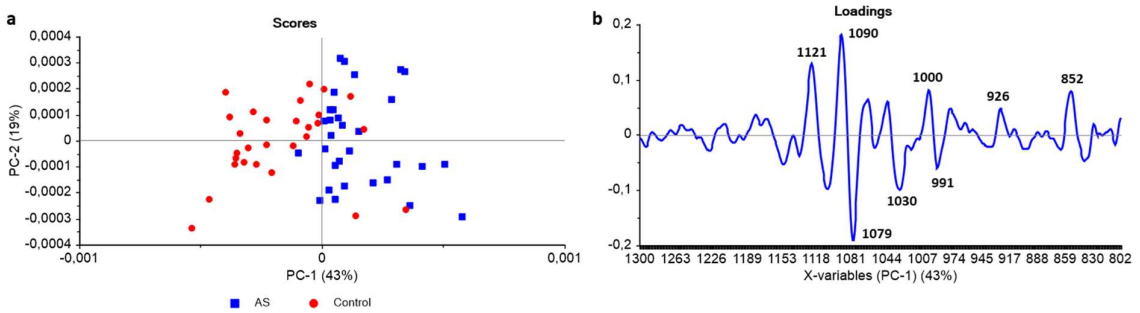


Figure 3. PCA a) scores and b) loading plots for FTIR spectra of control and AS groups in the 1300-800 cm^{-1} spectral region.

Based on the variance patterns revealed through PCA, we developed supervised classification models utilizing PCA-LDA and linear SVM, which incorporated the best spectral region (1300–800 cm^{-1}) determined from the PCA loading plots. These supervised models provided a robust framework for discriminating between AS and control samples, highlighting the potential of this combined multivariate analysis approach for non-invasive diagnostics and enhanced comprehension of the biochemical alterations associated with atherosclerosis.

High classification performances, with accuracies exceeding 88%, were achieved for the 1800-800 cm^{-1} region. Over 70% accuracy was obtained for the entire spectral region (4000-450 cm^{-1}), C-H region (3020-2800 cm^{-1}) and the protein region (1700-1500 cm^{-1}) using LDA. The fingerprint region (1300-800 cm^{-1}) yielded the highest classification performance with LDA classifiers achieving perfect scores of 96.61% accuracy. Accuracy is defined as the ratio of correctly classified samples to the total, expressed as a percentage. The discrimination plot and confusion table of Mahalanobis classifier model is shown in Fig. 4. This model achieved 96.61% accuracy and 93.5% true positive rate (sensitivity), for the AS class. The specificity (true negative rate) was 100%. For the LDA classifier applied to the fingerprint region, all of the 28 control samples and 29 out of 31 AS samples were correctly classified (Fig. 4, Table S1). These results highlight the robustness of the fingerprint region for differentiating AS and control samples, underscoring its potential as a reliable diagnostic tool.

The classification performance of the linear SVM model using the 1800-800 cm^{-1} spectral region is shown in Fig. 5. This model achieved a high overall accuracy of 93.22%, with 100% sensitivity and 85.71% specificity, correctly classifying all 31 AS samples and 24 out of 28 control samples (Fig. 5, Table S2). The features extracted from this region offer strong discriminative power for differentiating AS patients from controls, based on the high classification performance. The excellent performance of the linear kernel compared to the

ARTICLE IN PRESS

radial basis function (RBF) kernel indicates that the data are likely to be linearly separable in the selected feature space, which simplifies the model and improves interpretability. However, the explained variances for validation scores were slightly lower than those of the calibration scores, potentially indicating limitations associated with the relatively small dataset size. Furthermore, the model achieved perfect sensitivity but slighty lower specificity, demonstrating a robust ability to identify AS samples despite a tendency to misclassify some control subjects. This might be attributed to either class imbalance or overlapping spectral patterns in the control group. In addition, linear SVM model using the 1300–800 cm^{-1} region also showed strong classification performance (91.53% accuracy, 100% sensitivity, and 82% specificity). The corresponding classification plot and confusion matrix for this model are provided in supplementary material Fig. S3 and Table S3. A table summarizing the model type, number of PCs (for LDA), C-values (for SVM), and performance metrics (accuracy, sensitivity, specificity) for both calibration and validation is included in the supplementary material as Table S4. Moreover, regarding model validation, we also performed analyses using the training-test split approach. However, validation accuracies were significantly lower compared to training on the entire dataset (Fig. S4, Table S5).

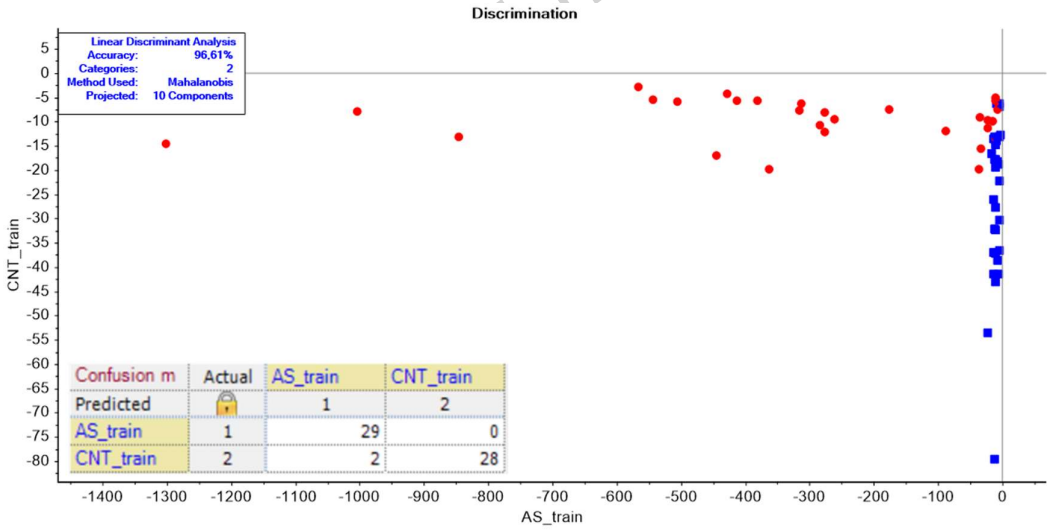
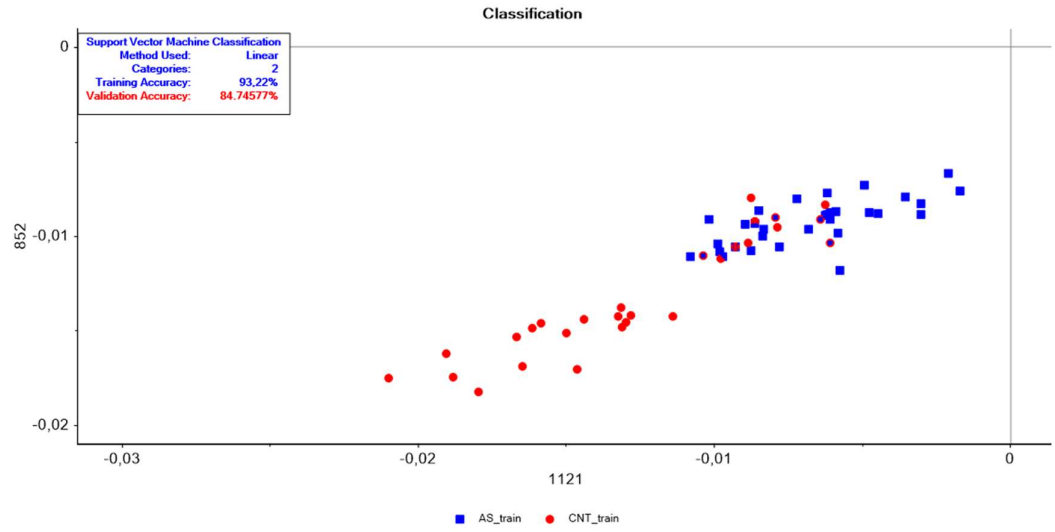


Figure 4. Discrimination plot and confusion matrix for serum samples from the AS and control group in the 1300-800 cm^{-1} region are shown. Control (CNT) and atherosclerosis (AS) are compared.

ARTICLE IN PRESS



Confusion matrix	Actual	AS_train	CNT_train
Predicted		1	2
AS_train	1	31	4
CNT_train	2	0	24

Figure 5. SVM classification plot and confusion matrix for serum samples from the AS and CNT group in the 1800-800 cm^{-1} region.

ARTICLE IN PRESS

Supervised multivariate analysis methods have drawn a lot of interest and use in the diagnosis of a variety of diseases. These algorithms were used as a holistic strategy to create a quick and accurate diagnosis tool, focusing on many fingerprint spectral windows that encompass a range of spectral variables possibly linked to the onset of the disease. The best discrimination was achieved in the 1300-800 cm^{-1} region, with 96.61% accuracy, 93.5% sensitivity, and 100% specificity values using discriminant analysis. The higher accuracy observed in AS classification for the 1300-800 cm^{-1} region, where functional groups specific to molecules such as DNA and RNA are located, suggests a potential association of these molecules with AS pathogenesis. Although the classification performances for the 1800-800 cm^{-1} region were above 88% accuracy, and over 70% accuracy was achived for entire spectral region (4000-450 cm^{-1}), C-H region (3020-2800 cm^{-1}) and the protein region (1700-1500 cm^{-1}) using LDA, these were not as high as those for the fingerprint region. The lower accuracy and sensitivity in these regions may reflect the greater contribution of disease and patient-related factors, such as subtypes, stages, comorbidities and nutritional status, to the serum lipid or protein profile. The presence of comorbidities such as diabetes mellitus, hypercholesterolemia and the presence of coronary artery disease (CAD) in their family history of AS and control groups are reported in previous study.[33] Comorbidities such as diabetes and

ARTICLE IN PRESS

hyperlipidemia were recorded for all individuals and were relatively balanced between AS and control groups, minimizing their impact on spectral discrimination.[33]

One limitation of this study is the relatively small sample size. To verify the sample size, achieved power analysis has been performed by G*Power software for the spectral calculations and over 99% power values were obtained. Our achieved power values imply that the sample size in each group is sufficient to draw conclusions from our findings. Nevertheless, future studies should involve larger cohorts to validate these findings and assess the applicability of FTIR-based diagnostics across different stages and subtypes of atherosclerosis. Moreover, the current study presents compelling evidence of FTIR's diagnostic capabilities. However, its generalizability across diverse populations and comorbid conditions requires further investigation. Additionally, examining the impact of dietary and lifestyle interventions on the identified spectral biomarkers could enhance the clinical relevance of this approach.

Conclusion

This study highlights the potential of FTIR spectroscopy, coupled with advanced multivariate analysis, as a powerful diagnostic tool for detecting subtle molecular changes in atherosclerosis. Compared to conventional imaging modalities such as CT angiography or MRI, FTIR offers rapid, low-cost, and molecular-level diagnostic capabilities. Although direct comparison is constrained due to differing mechanisms, ATR-FTIR spectroscopy may serve as a complementary and non-invasive screening tool with high sensitivity. Our findings confirm significant alterations in lipid metabolism, protein structure, and nucleic acid content and conformation in patients with atherosclerosis, providing deeper insights into the disease's pathophysiology. The integration of FTIR spectroscopy with machine learning algorithms holds great promise for developing personalized diagnostic and prognostic models, enabling early intervention and improving patient outcomes. With further refinement, FTIR spectroscopy could become a routine clinical tool for early screening, risk assessment, and monitoring of atherosclerosis progression. Its ability to detect small molecular changes in the blood, combined with machine learning algorithms, could revolutionize cardiovascular disease management by enabling diagnosis long before symptoms appear.

Material and Methods

Patients and sample collection. All methods, experimental protocols and experiments were approved by the Yuksek Ihtisas University Ethical Committee with the file number 2023/02/03 in accordance with the requirements of Helsinki Declaration. Informed consent of participants

ARTICLE IN PRESS

was obtained following the requirements of the research ethics boards and all associated methods were conducted in accordance with approved guidelines and regulations for human experimental research.

Whole blood was collected in EDTA containing tubes and samples were immediately centrifuged at 1900 g for 10 min, supernatant was again centrifuged for 10 min at 10.000 g at 4 °C and serum samples were transferred into RNase/DNase free tubes and stored at − 80 °C for future experiments.[33]

The samples were collected from 31 patients with atherosclerosis (AS, mean age 63.5) and 28 control individuals without a history of cardiovascular disease (C, mean age 55.3). There was no significant difference in clinical or demographic characteristics of the patients and control groups including age, LDL cholesterol, HDL cholesterol or blood glucose levels (Table 1).

Table 1. Demographic and clinical parameters of blood samples are given together with p values calculated by student's t-test.

	AS (n=31)	Control (n=28)	p-value
Age (mean± SD)	63.5± 8.4	55.3± 8.5	0.08
LDL (mean ± SD)	120.1± 46.4	117.5± 27	0.81
HDL (mean± SD)	39.8± 7.6	43.2±6.6	0.09
Triglyceride (mean± SD)	162.4± 91.6	144.3± 44.4	0.35
Glucose (mean± SD)	125.6± 54.4	112.6± 41.6	0.32

ATR-FTIR spectroscopy. Frozen serum samples were thawed and mixed at room temperature prior to spectroscopic analysis. Freezing and storing the samples at -80°C preserves the stability of biomolecular analytes over time and prevents degradation due to enzymatic activity, ensuring reliable and reproducible spectral results. Afterwards they were directly put on the ATR crystal and scanned. The acquisition of ATR-FTIR spectra of all samples were collected by Perkin Elmer Frontier model FTIR spectrometer (Perkin Elmer Inc., USA) equipped with a Quest single reflection ATR accessory (Specac Ltd., UK). 1 µl of serum samples were placed on a diamond crystal plate and dried with a mild nitrogen gas flux for 2 min to remove excess unbound water. IR spectra of the samples were scanned in the 4000-450 cm^{-1} wavenumber region and 64 scans were taken for each interferogram with 4 cm^{-1} spectral resolution at room temperature.[7] In order to eliminate the effects of water molecules in the air,

ARTICLE IN PRESS

the spectrum of air prior to sample spectra acquisition was recorded as a background under identical conditions with the samples and subtracted automatically from all the spectra. At least three randomly taken samples (1 µl of serum each) from each sample were prepared and then scanned to check the reproducibility of the identical spectra. The average spectra of these replicates were used in further data analysis. These average spectra were baseline corrected and then normalized with respect to the amide A band for visual demonstration. The band positions were determined using the center of mass of the width at 75% of band intensity.[15] Band ratio analysis was performed as described in our previous work.[7] Briefly, specific spectral bands were integrated using Perkin Elmer Spectrum software (v10.03.06) following baseline correction, and relative area ratios were calculated for semi-quantitative comparison. For full methodological details, please refer to Section 2.2 and Table 3 of our previously published work.[7]

Multivariate analysis methods. Principal component analysis (PCA), an unsupervised multivariate analysis method, was performed to analyze the discrimination of the diseased group (AS) from controls based on spectral differences. PCA simplifies the evaluation of extensive datasets by transforming a large set of variables into a smaller set of principal components (PCs), which retain most of the variance in the original dataset. This transformation facilitates the visualization and identification of distinct sample classes using two key parameters; eigen vectors and eigen values. The eigen values indicate the variation in each principle component, and the eigenvectors (loadings) provide the weighting function for obtaining the PCA scores.[7,15,60] In the present study, PCA was applied to baseline corrected, vector-normalized second derivative spectra within the 4000-450 cm^{-1} (whole) region and specific sub-regions, including 3020-2800 cm^{-1} (lipid), 1700-1500 cm^{-1} (protein), 1300-800 cm^{-1} (fingerprint) by using The Unscrambler X, Version 10.3 program (CAMO Software Inc., Oslo, Norway). PCA was performed using the Singular Value Decomposition (SVD) algorithm, with mean-centered data and with full cross-validation (known as leave-one-out cross-validation (LOOCV)). PCA results are given by score and loading plots.[15]

Supervised classification techniques provide the identification of local models for studied groups and the prediction of probable class memberships. For this purpose, we used discriminant analysis and support vector machine (SVM) classification.

Discriminant analysis is employed to develop classification rules for pre-specified classes, facilitating the assignment of unknown samples to the most probable class and enabling the interpretation of intergroup differences. The optimal parameters for classifying the samples are

ARTICLE IN PRESS

determined, and the developed model is subsequently used to classify unknown samples. Linear discriminant analysis (LDA) algorithm utilizes Mahalanobis distance calculation between the samples of each class, which can include or exclude Bayesian probability terms. LDA relies on the assumption of a normal distribution and the assumption that the covariance matrices of the two groups are similar, implying that the variability within each group possesses an identical structure. Quadratic Discriminant Analysis (QDA) algorithm is also based on a Mahalanobis distance calculation between the samples for each class, utilizing Bayesian probability terms to correct for classes having different sizes. In contrast to LDA, QDA forms an individual variance model for each class, employing a distinct variance-covariance matrix for every class. Consequently, QDA performs better than LDA when analyzing classes with varying within-category variances. However, both LDA and QDA are parametric approaches that are highly influenced by ill-conditioned data; these issues can be solved by the application of PCA to the preprocessed spectral data before LDA or QDA. The key disadvantages of QDA are that it performs worse than LDA for small-sized datasets and has a higher risk of overfitting than LDA. We chose Mahalanobis classification because distance-based classification typically offers greater stability in high-dimensional spectral data. It remains unaffected by irrelevant spectral variation and helps avoid the overfitting issues often encountered with unconstrained QDA.[61] Thus, we can achieve superior accuracy and reproducibility in our IR dataset.

The inputs to the classification that processed using The Unscrambler X 10.3 (CAMO Software AS, Norway) software were the vector normalized second derivative spectra in the 1300-800 cm^{-1} range, with known class labels. Initially, the category variable column was incorporated into the data matrix, and spectra from all sample classes were subsequently used to construct a training set. The classification was based on a Mahalanobis distance classifier, utilizing the projections of the first 10 principal components, which were selected based on the amount of explained variance (≥95%) and cross-validation performance. Prior probabilities were calculated from the training set. The results are presented in the form of discrimination plots, prediction and confusion matrices. The prediction matrix exhibits the predicted class for each sample along with the probability of membership for each class. In contrast, the confusion matrix reports both the predicted and actual classifications of samples. Each row in the confusion matrix represents instances within a predicted class, while each column corresponds to instances within an actual class.[7,9,62]

The support vector machine (SVM) classifier operates by identifying the optimal hyperplane that best separates data points belonging to different classes in a high-dimensional space. When data is not linearly separable, SVM utilizes kernel functions (such as radial basis function,

ARTICLE IN PRESS

polynomial, or linear) to transform the input space into a higher-dimensional feature space, enabling the construction of a separating hyperplane. The decision boundary is defined by a subset of training data points, known as support vectors, which are essential in determining the position and orientation of the hyperplane. Model performance was optimized by adjusting the ν hyperparameter of the ν-SVM, which controls the trade-off between training error and model complexity. Both linear and RBF kernels were tested, and the linear kernel yielded better performance with ν value of 0.5.[7,61,63,64] In this study, linear SVM was implemented using the Unscrambler X, Version 10.3 program (CAMO Software Inc., Oslo, Norway), and model parameters were chosen using cross-validation to ensure generalizability.

Statistical analysis. Statistical analyses were performed using an unpaired (independent) t-test in GraphPad Prism 6 (GraphPad Software, Inc., USA) to compare the means of two independent groups and determine statistical significance. A p-value < 0.05 was considered statistically significant. Results were expressed as mean ± SEM. Statistical significance levels were denoted as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ and $****p < 0.0001$.

ARTICLE IN PRESS

References

1. Ross, R. The pathogenesis of atherosclerosis: A perspective for the 1990s. *Nature* **362**, 801–809 (1993).
2. Cagnina, A. *et al.* Atherosclerosis, an inflammatory disease. *Rev. Med. Liege* **77**, 302–309 (2022).
3. Rafieian-Kopaei, M., Setorki, M., Doudi, M., Baradaran, A. & Nasri, H. Atherosclerosis: Process, indicators, risk factors and new hopes. *International Journal of Preventive Medicine* (2014).
4. Syed, M. B. J. *et al.* Emerging techniques in atherosclerosis imaging. *Br. J. Radiol.* **92**, 20180309 (2019).
5. Martin, S. S. *et al.* 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data from the American Heart Association. *Circulation* **149**, E347–E913 (2024).
6. Naumann, D., Fabian, H. & Lasch, P. FTIR Spectroscopy of Cells, Tissues and Body Fluids. in *Biological and Biomedical Infrared Spectroscopy* (eds. Barth, A. & Haris, P. I.) 312–354 (IOS Press, 2009). doi:10.3233/978-1-60750-045-2-312.
7. Yonar, D. *et al.* Rapid diagnosis of malignant pleural mesothelioma and its discrimination from lung cancer and benign exudative effusions using blood serum. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1868**, (2022).
8. Byrne, H. J., Bonnier, F., McIntyre, J. & Parachalil, D. R. Quantitative analysis of human blood serum using vibrational spectroscopy. *Clin. Spectrosc.* **2**, (2020).
9. Pang, W. *et al.* Serum-based ATR-FTIR spectroscopy combined with multivariate analysis for the diagnosis of pre-diabetes and diabetes. *Analyst* **149**, (2023).
10. Sala, A. *et al.* Rapid analysis of disease state in liquid human serum combining infrared spectroscopy and “digital drying”. *J. Biophotonics* **13**, (2020).
11. Ollesch, J. *et al.* It’s in your blood: spectral biomarker candidates for urinary bladder

ARTICLE IN PRESS

cancer from automated FTIR spectroscopy. *J. Biophotonics* **7**, 210–221 (2014).
12. Backhaus, J. *et al.* Diagnosis of breast cancer with infrared spectroscopy from serum samples. *Vib. Spectrosc.* **52**, 173–177 (2010).
13. Nogueira, M. S. *et al.* FTIR spectroscopy as a point of care diagnostic tool for diabetes and periodontitis: A saliva analysis approach. *Photodiagnosis Photodyn. Ther.* **40**, (2022).
14. Contreras-Rozo, J. A., Mata-Miranda, M. M., Vazquez-Zapien, G. J. & Delgado-Macuil, R. J. Infrared spectroscopy technique: An alternative technology for diabetes diagnosis. *Biomedical Signal Processing and Control* vol. 86 (2023).
15. Yonar, D., Ocek, L., Tiftikcioglu, B. I., Zorlu, Y. & Severcan, F. Relapsing-Remitting Multiple Sclerosis diagnosis from cerebrospinal fluids via Fourier transform infrared spectroscopy coupled with multivariate analysis. *Sci. Rep.* **8**, 1025 (2018).
16. Wang, X. *et al.* A study of Parkinson's disease patients' serum using FTIR spectroscopy. *Infrared Phys. Technol.* **106**, (2020).
17. Ami, D., Mereghetti, P. & Maria, S. Multivariate Analysis for Fourier Transform Infrared Spectra of Complex Biological Systems and Processes. in *Multivariate Analysis in Management, Engineering and the Sciences* (2013). doi:10.5772/53850.
18. Houhou, R. & Bocklitz, T. Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. *Analytical Science Advances* vol. 2 (2021).
19. Tosi, G. *et al.* Vibrational spectroscopy as a supporting technique in clinical diagnosis and prognosis of atherosclerotic carotid plaques. *Anal. Quant. Cytol. Histol.* **34**, (2012).
20. Barrett, H. E., Mulvihill, J. J., Cunnane, E. M. & Walsh, M. T. Characterising human atherosclerotic carotid plaque tissue composition and morphology using combined spectroscopic and imaging modalities. *Biomed. Eng. Online* **14**, (2015).
21. Li, C. *et al.* Biochemical characterization of atherosclerotic plaque constituents using FTIR spectroscopy and histology. in *Journal of Biomedical Materials Research - Part A* vol. 64 (2003).
22. Anastassopoulou, J. *et al.* Infrared Spectroscopic Study and Mathematical Simulations of Carotid Atherosclerosis. *In Vivo (Brooklyn).* **36**, (2022).
23. Mamarelis, I. *et al.* Oxidative stress and atherogenesis. An FT-IR spectroscopic study. *In Vivo (Brooklyn).* **24**, (2010).
24. Marzec, K. M. *et al.* Vascular diseases investigated ex vivo by using Raman, FT-IR and complementary methods. *Pharmacological Reports* vol. 67 (2015).
25. Peters, A. S. *et al.* Serum-infrared spectroscopy is suitable for diagnosis of atherosclerosis and its clinical manifestations. *Vib. Spectrosc.* **92**, (2017).
26. Kęsik JJ, Paja W, Jakubczyk P, Khalavka M, Terlecki P, Iłżecki M, Rzad W, D. J. Determination of spectroscopy marker of atherosclerotic carotid stenosis using FTIR-ATR combined with machine learning and chemometrics analyses. *Nanomedicine Nanotechnology, Biol. Med.* **62**, (2024).
27. Ferdous, B. *et al.* Serum Lipid Profile in Ischemic Heart Disease: A Cross-Sectional Analytical Study. *Bangladesh J. Med. Biochem.* **7**, (2014).
28. Anroedh, S. *et al.* Plasma concentrations of molecular lipid species predict long-term clinical outcome in coronary artery disease patients. *J. Lipid Res.* **59**, (2018).
29. Zakiev, E. R. *et al.* Lipid composition of circulating multiple-modified low density lipoprotein. *Lipids Health Dis.* **15**, (2016).
30. Siri-Tarino, P. W., Sun, Q., Hu, F. B. & Krauss, R. M. Saturated fat, carbohydrate, and cardiovascular disease. *American Journal of Clinical Nutrition* vol. 91 (2010).
31. Lu, Y. *et al.* Microbiota-derived short-chain fatty acids: Implications for cardiovascular and metabolic disease. *Front. Cardiovasc. Med.* **9**, 1–17 (2022).
32. Pielesz, A., Biniaś, D., Waksmańska, W. & Bobiński, R. Lipid bands of approx. 1740

ARTICLE IN PRESS

cm−1 as spectral biomarkers and image of tissue oxidative stress. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* vol. 286 (2023).
33. Telkoparan-Akillilar, P. & Cevik, D. Identification of miR-17, miR-21, miR-27a, miR-106b and miR-222 as endoplasmic reticulum stress-related potential biomarkers in circulation of patients with atherosclerosis. *Mol. Biol. Rep.* **48**, 3503–3513 (2021).
34. Severcan, F. *et al.* Decoding myasthenia gravis: advanced diagnosis with infrared spectroscopy and machine learning. *Sci. Rep.* **14**, 1–16 (2024).
35. Kotoulas, E. *et al.* The influence of diabetes on atherosclerosis and amyloid fibril formation of coronary arteries. A FT-IR spectroscopic study. *Hell. J. of Atheroscler.* **8**, (2018).
36. Ursini, F., Davies, K. J. A., Maiorino, M., Parasassi, T. & Sevanian, A. Atherosclerosis: Another protein misfolding disease? *Trends in Molecular Medicine* vol. 8 (2002).
37. Fert-Bober, J., Holewinski, R. J., Kooren, J. A. & Van Eyk, J. E. Posttranslational Modifications. in *Atherosclerosis* 437–451 (2015). doi:https://doi.org/10.1002/9781118828533.ch34.
38. Elliott, K. J. & Eguchi, S. Phosphorylation Regulation by Kinases and Phosphatases in Atherosclerosis. in *Atherosclerosis* 453–465 (2015). doi:https://doi.org/10.1002/9781118828533.ch35.
39. Li, J. *et al.* Progress of Fluorescent Probes for Protein Phosphorylation and Glycosylation in Atherosclerosis. *Chem. - A Eur. J.* **30**, 1–9 (2024).
40. Triska, J., Mathew, C., Zhao, Y., Chen, Y. E. & Birnbaum, Y. Circular RNA as Therapeutic Targets in Atherosclerosis: Are We Running in Circles? *J. Clin. Med.* **12**, (2023).
41. Meiliana, A. & Wijaya, A. MicroRNAs in Lipid Metabolism and Atherosclerosis. *Indones. Biomed. J.* **6**, 3 (2014).
42. Churov, A., Summerhill, V., Grechko, A., Orekhova, V. & Orekhov, A. MicroRNAs as potential biomarkers in atherosclerosis. *Int. J. Mol. Sci.* **20**, (2019).
43. Chen, W. J., Yin, K., Zhao, G. J., Fu, Y. C. & Tang, C. K. The magic and mystery of MicroRNA-27 in atherosclerosis. *Atherosclerosis* **222**, 314–323 (2012).
44. Madrigal-Matute, J., Rotllan, N., Aranda, J. F. & Fernández-Hernando, C. MicroRNAs and atherosclerosis. *Curr. Atheroscler. Rep.* **15**, (2013).
45. Raitoharju, E. *et al.* MiR-21, miR-210, miR-34a, and miR-146a/b are up-regulated in human atherosclerotic plaques in the Tampere Vascular Study. *Atherosclerosis* **219**, 211–217 (2011).
46. Fichtlscherer, S. *et al.* Circulating microRNAs in patients with coronary artery disease. *Circ. Res.* **107**, 677–684 (2010).
47. Ren, J. *et al.* Signature of circulating MicroRNAs As potential biomarkers in vulnerable coronary artery disease. *PLoS One* **8**, (2013).
48. Telkoparan-Akillilar, P., Cevik, D. & Yilmaz, A. Expression patterns of miR-34a, miR-125b, miR-221 and antioxidant gene NRF2 in plasma samples of patients with atherosclerosis. *J. Biosci.* **47**, 1–12 (2022).
49. Temiz, N. A., Donohue, D. E., Bacolla, A., Luke, B. T. & Collins, J. R. The Role of Methylation in the Intrinsic Dynamics of B- and Z-DNA. *PLoS One* **7**, e35558 (2012).
50. Dai, Y., Chen, D. & Xu, T. DNA Methylation Aberrant in Atherosclerosis. *Front. Pharmacol.* **13**, 1–12 (2022).
51. Shah, N. R. & Mahmoudi, M. The role of DNA damage and repair in atherosclerosis: A review. *Journal of Molecular and Cellular Cardiology* vol. 86 (2015).
52. Ishida, M., Sakai, C. & Ishida, T. Role of DNA damage in the pathogenesis of atherosclerosis. *Journal of Cardiology* vol. 81 (2023).

ARTICLE IN PRESS

ARTICLE IN PRESS

53. Petibois, C. *et al.* Determination of Glucose in Dried Serum Samples by Fourier-Transform Infrared Spectroscopy. *Clin. Chem.* **45**, (1999).
54. Poznyak, A. *et al.* The Diabetes Mellitus–Atherosclerosis Connection: The Role of Lipid and Glucose Metabolism and Chronic Inflammation. *International Journal of Molecular Sciences* vol. 21 (2020).
55. Ye, J. *et al.* Diabetes Mellitus Promotes the Development of Atherosclerosis: The Role of NLRP3. *Front. Immunol.* **13**, (2022).
56. Aronson, D. & Rayfield, E. J. How hyperglycemia promotes atherosclerosis: molecular mechanisms. *Cardiovasc. Diabetol.* **1**, 1 (2002).
57. Mastrogiacomo, L. *et al.* The Effects of Hyperglycemia on Early Endothelial Activation and the Initiation of Atherosclerosis. *Am. J. Pathol.* **193**, 121–133 (2023).
58. Flynn, M. C. *et al.* Transient Intermittent Hyperglycemia Accelerates Atherosclerosis by Promoting Myelopoiesis. *Circ. Res.* **127**, 877–892 (2020).
59. Yuan, T. *et al.* New insights into oxidative stress and inflammation during diabetes mellitus-accelerated atherosclerosis. *Redox Biol.* **20**, 247–260 (2019).
60. Wang, L. & Mizaikoff, B. Application of multivariate data-analysis techniques to biomedical diagnostics based on mid-infrared spectroscopy. *Analytical and Bioanalytical Chemistry* vol. 391 (2008).
61. Morais, C. L. M., Lima, K. M. G., Singh, M. & Martin, F. L. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* vol. 15 (2020).
62. Ollesch, J. *et al.* FTIR spectroscopy of biofluids revisited: an automated approach to spectral biomarker identification. *Analyst* **138**, 4092 (2013).
63. Gholami, R. & Fakhari, N. Support Vector Machine: Principles, Parameters, and Applications. in *Handbook of Neural Computation* (2017). doi:10.1016/B978-0-12-811318-9.00027-2.
64. Chatchawal, P. *et al.* Detection of human cholangiocarcinoma markers in serum using infrared spectroscopy. *Cancers (Basel).* **13**, (2021).

ARTICLE IN PRESS

Acknowledgments

The authors would like to thank Prof. Dr. Mustafa PAC for his support during the collection of patient samples.

Author contributions

All authors contributed to the study conception and design. DC, PTA and DY conceptualized the study. DC, PTA and DY performed the experiments. DY performed data analysis, and data interpretation. DC and DY wrote the initial manuscript. All authors corrected and approved the final manuscript.

Funding

This research did not receive support from any funding organization.

ARTICLE IN PRESS

Data Availability The data utilized and/or analyzed in the current study is available from the corresponding author upon reasonable request.

Declarations

Conflict of interests The authors declare no competing interests.

Ethical Approval The study was approved by the Ethics Committee of Yuksek Ihtisas University (approval number: 2023/02/03).

ARTICLE IN PRESS