Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

MACHINE LEARNING COMBINED WITH INFRARED SPECTROSCOPY FOR DETECTION OF HYPERTENSION PREGNANCY: TOWARDS FETAL AND PREGNANT BLOOD ANALYSIS.

Sara Maria Santos Dias da Silva[1], Marcelo Saito Nogueira[2], Jaqueline Maria Brandão Rizzato[1], Simone de Lima Silva[1], Sheila Cavalca Cortelli[1], Roger Borges[3], Herculano da Silva Martinho[3], Rodrigo Augusto Silva[1], and Luis Felipe das Chagas e Silva de Carvalho[1]

[1]Universidade de Taubate
[2]Tyndall National Institute
[3]Universidade Federal do ABC Centro de Ciencias Naturais e Humanas

May 05, 2024

Abstract

Biochemical changes in the cervix during labor are not well understood, in part because of a lack of technology capable of safely probing the pregnant cervix in vivo. FT-IR spectroscopy has the potential to address these needs because it is a non-invasive optical technique that can sensitively detect changes in biochemical components. A total of 30 pregnant participants undergoing either spontaneous or induced labor were recruited. We detected several biochemical changes during labor, including a significant decrease in FT-IR spectral features associated with collagen and other extracellular matrix proteins attributed to collagen dispersion, an increase in spectral features associated with blood, and in features indicative of lipid-based molecules. Our results have demonstrated that FT-IR spectroscopy is sensitive to multiple biochemical remodeling changes in the cervix during labor. FT-IR spectroscopy may be a valuable noninvasive tool for objective cervical assessment to potentially guide clinical labor management.

INTRODUCTION

Hypertensive disorders of pregnancy (HDP, such as pre-eclampsia and eclampsia) are the second leading cause of maternal mortality worldwide and a leading cause of preterm birth (1,2,3). These conditions are associated with fetal and neonatal mortality and account for 14.1% of maternal deaths (1). Pre-eclampsia (PE) and eclampsia complicate between 3-5% of pregnancies (3,4). In Brazil, PE is also a leading cause of perinatal mortality. PE significantly compromises maternal health and has serious consequences for the fetus and newborn (5). A systematic review found an incidence of 1.5% for PE and 0.6% for eclampsia (6). In more developed areas of Brazil, it is estimated to be 0.2%, with 0.8% of mothers dying from it, while in disadvantaged regions it is estimated to be 8.1%, with 22.0% of mothers dying from it (7).

Considering the greater flexibility for diagnosis, and according to the Brazilian Federation of Gynecology and Obstetrics Associations (FEBRASGO), evaluations are considered adequate in isolated urine samples with a proteinuria/creatinuria ratio (both in mg/dl) equal to or greater than 0.3. In the absence of these diagnostic capabilities, a proteinuria with at least a 1+ in the dipstick may be considered as soon as the quality of the method has been assured. The intensity of proteinuria should no longer be the sole determinant of maternal prognosis or decision making (5). PE is caused by arterial hypertension (AH), first observed

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

after 20 weeks, in addition to proteinuria, and may be superimposed on another hypertensive state (8). Thus, in the absence of proteinuria, the diagnosis of PE may be based on the presence of headache, blurred vision, abdominal pain, or abnormal laboratory tests such as thrombocytopenia (less than 100,000/mm^3), elevated liver enzymes (twice the baseline), renal impairment (above 1.1 mg/dl or twice the baseline), or even pulmonary edema and visual or cerebral disturbances such as headache, scotomas, or seizures (5).

The diagnosis of PE should be anticipated in pregnant women who develop AH and significant proteinuria after the 20th week of gestation (except for hydatidiform mole, in which PE may occur prior to the 20th week of gestation). If the rise in blood pressure and the proteinuria occurs after 20 weeks of gestation in a primigravida with a family history (especially a sister or mother) of pre-eclampsia or eclampsia, the chance of a correct diagnosis of pre-eclampsia is greater than 90% (5). In these patients, special attention should be given during prenatal care to diagnose pre-eclampsia as early as possible (5,9). This assessment should be limited to measurement of platelets, creatinine, uric acid, and baseline proteinuria (e.g., proteinuria/creatinine ratio in a urine specimen). Accurate determination of gestational age by ultrasound is essential in the first trimester of pregnancy. Doppler evaluation of the uterine arteries after 23 weeks' gestation can assess the presence or absence of adequate placental implantation by normal resistance indices (10,11).

Treatment management includes preconception counseling, perinatal blood pressure control and management of associated complications, timely fetal delivery, and postpartum monitoring. New strategies to treat the clinical signs of pre-eclampsia and to prolong pregnancy are still under investigation (12). Nevertheless, the management of PE requires its early detection. There are many publications in the literature suggesting methods to detect the risk of PE, according to FEBRASGO (5). Among the various options, the use of uterine arterial Doppler is noteworthy, and the detection of plasma components, such as proteins of placental origin or due to angiogenic dysregulation, is performed in the first or second trimester and shows limited accuracy, in addition to the difficulty of ensuring the qualification of its measurement (12,13).

FT-IR spectroscopy provides characterization of molecular bonds and functional groups present in solid, liquid, and gaseous samples. For example, qualitative and semi-quantitative information on the biochemical components of saliva components can be obtained in less than 5 minutes (14). Over the years, vibrational spectroscopy has shown tremendous potential in the diagnosis of disease (14-16) and in the identification of molecular components in biological tissues (16,17) and in biofluids (14,15,18,19). One of the most widely used vibrational spectroscopy techniques is Fourier Transform Infrared (FTIR) spectroscopy due to its cost effectiveness and technological maturity, which has led to portable high-end detectors that can potentially be used in the clinic. In the last few decades, FTIR spectroscopy has been used for the development of new alternative screening and diagnostic methods based on body fluids (19,20). We hypothesized that blood plasma biomarkers could potentially contain angiogenic factors, which have emerged as important biomarkers in pre-eclampsia, since the imbalance of angiogenic markers is central to the pathogenesis of PE (21).

Despite the large number of predictive factors, the use of plasma markers related to angiogenesis/antiangiogenesis imbalance has been presented in the literature as a promising tool for the early detection of PE. The evaluation of biomarkers for PE has been the subject of numerous studies and may be useful in the early diagnosis of PE. Ideally, the biomarker should be easy to perform, inexpensive, and allow detection as early as possible, preferably in the first trimester of pregnancy, before hypertension develops (13,22); plasma may contain other biomarkers associated with predisposing factors to PE, including gestational diabetes mellitus, renal disease, increased placental mass as seen in multiparity, molar pregnancy, and primigravida, previous or family history of PE, chronic hypertension, collagenases, black race, obesity, and thrombophilia (5,7).

Vibrational spectroscopy is one of the next generation technologies for real-time and multiplexed biomarker detection. Optical spectroscopy is a versatile, rapid, and non-destructive analytical technique for the analysis of biological tissues and biofluids. Fourier Transform Infrared (FT-IR) spectroscopy is based on the absorption of radiation by sample molecules, providing information about their structure and composition (15). Most of the molecular-specific information is concentrated in the fingerprint region between 600 and 1800 cm-1, which provides valuable information to clinical/medical professionals via optical biopsy (i.e., diagnosis of

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

lesions and pathologies characterized by molecular-specific biomarkers present in human tissues, cells and secretions) (23).

In our study, we have evaluated the feasibility of detecting PE biomarkers based on FTIR spectra of blood plasma samples. Our analysis was based on wavenumbers selected by partial least-squares discriminant analysis (PLS-DA) on the spectral fingerprint region. This selection prioritized the classification of plasma samples into control and PE study groups with maximum accuracy. The corresponding vibrational modes associated with selected wavenumbers defined the biomarkers potentially associated with the development of PE. Corresponding modes associated with selected wavenumbers defined the biomarkers potentially associated with PE development. The feasibility of PE detection was evaluated by the classification accuracy. We believe that our study can improve the early diagnosis. It has been observed that oxidative stress induces apoptosis and deterioration, promotes the accumulation of reactive oxygen species and the generation of free radicals, and increases the leakage of apoptotic aggregates into the maternal circulation, resulting in systemic endothelial dysfunction (24,25).

Thus, the differentiation of spectra can provide new information for a possible new screening protocol for risk assessment through biomarkers. Countries with lower rates of development have a greater chance of the population developing risks in the face of pregnancy, in Brazil about 23% of maternal deaths are caused by hypertensive conditions related to arterial hypertension. Therefore, evaluating biomarkers that may predispose to preeclampsia is extremely important for health professionals to prevent and thus reduce the number of conditions that develop and lead to the death of pregnant women (26).

Materials and methods

Clinical protocol

A total of 30 pregnant women and 30 related newborns admitted to the Municipal University Hospital of Taubaté (HMUT) for induction of labor or spontaneous labor were recruited and enrolled in a retrospective observational pilot study under a protocol approved by the Research Ethics Committee (CEP) of the University of Taubaté (approval number CAAE: 34126120.9.0000.5501) according to the guidelines of the Declaration of Helsinki. All study participants were over 18 years of age. (Table 1)

Study groups : plasma isolated from blood from Control Pregnant Women (CG), consisting of 15 normotensive patients without pre-eclampsia; of Pregnant Women with Pre-eclampsia (GPré), consisting of 15 patients who developed pre-eclampsia (who were followed up in the HMUT high-risk prenatal care or with the diagnosis but in prenatal care at another location); 15 umbilical cords from newborn of pregnant women in the CG (FControl) and 15 umbilical cords from newborns of pregnant women in the GPré group (FPre).

Sample collection protocol.

Blood collection was performed at the HMUT, concurrently with routine collections during the patient's hospitalization for clinical follow-up in the obstetric pathology department, or at the time of delivery. The preferred site for vacuum blood collection was the antecubital fossa. The tube used was the EDTA-containing Vacuette® prepared to hold 4 mL of blood. After collection, the properly labeled tubes containing the blood were stored at 4°C until the sample was processed.

Inclusion and exclusion criteria

Inclusion - Pregnant women who had started prenatal care during the first half of pregnancy, gestational age (GA) [?] 20 weeks, with a single pregnancy and non-smokers, in whom at least one of the risk factors for the development of pre-eclampsia was present: obstetric history, nulliparity, primiparity, history of PE in a previous pregnancy; obesity (BMI[?] 30 kg/m2), age [?] 35 years; chronic arterial hypertension or without a risk factor present but with a diagnosis of PE at the time of peripartum. For the dating of pregnancy, gestational age was determined according to the date of the last menstrual period using Nagele's rule and/or corrected by ultrasound examination performed in the first half of pregnancy. In addition, the definition

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

of chronic arterial hypertension during pregnancy was based on the criteria of the American College of Gynecology and Obstetrics (5).

Exclusion - Pregnant women who presented during pregnancy conditions that could interfere with the interpretation of the biochemical values of interest or who did not reach the second half of pregnancy were excluded from the study; gestational diabetes; autoimmune diseases and miscarriages. Loss to follow-up was defined as voluntary withdrawal from the study and inability to access birth data.

Instrumentation and sample preparation

The collected tubes were processed at the Dental Research Center (CEPEO). For plasma collection, tubes containing maternal blood and cord blood were centrifuged at 3,000g for 10 minutes. The supernatant (plasma) was collected and stored at -80degC until use. EDTA was used as anticoagulant. The samples were collected in VACUETTE(r) Clot Activator Tubes. Samples were kept refrigerated at 4degC for approximately 3 hours until processed for plasma collection. The samples were stored at -80degC (6 months) until analyzed by FTIR.

FT-IR spectroscopy system setup

On a predetermined day for analysis, FT-IR spectra of saliva samples were collected using a Bruker Alpha II spectrometer equipped with an ATR-FTIR diamond crystal and heating functions to dry our samples. Our samples were pipetted without additives on the crystal of the spectrometer and dried for 1-3 minutes before data collection. Spectra from each sample were collected in triplicate. Once the spectra of one sample were collected, the crystal was sanitized with 70% alcohol and allowed to dry before the next sample was measured. This prevented cross-contaminating.

The FT-IR spectral fingerprint region between 600-1800 cm-1 was used to analyze the plasma samples. In the control group, two samples were analyzed, and six spectra were obtained. In the diabetic group with periodontitis, 12 samples were analyzed, and 36 spectra were obtained. Ten samples were analyzed, and 30 spectra were obtained in the diabetic group.

Data analysis

We evaluated heat maps based on the VIP (variable importance projection) score of each wavenumber of the FTIR spectra. To identify the wavelengths contributing most to the development of a potentially predictive PLS-DA model (Figure 2 and 3).

Moreover, as FTIR is based on vibrational bands made of a range of wavelenghts, the wavelenghts belonging to the sample band are clustered in the dendrogram, allowing their identification. In addition, bands that may be correlated to each other due to biochemistry basis of the pre-eclampsia or control groups can be linked together in the same node of the dendrogram.

The processing of the spectra was performed using RStudio software (27). The baseline correction was made using the least squares polynomial curve fitting method as described by Lieber and Mahadevan-Jansen. Partial least squares discriminant analysis (PLS-DA) was performed on the already corrected spectra using Pareto's scaling. More details have been described elsewhere (28). Receiver-characteristic operating curve (ROC) analysis was used to evaluate the diagnostic accuracy of a specific vibrational band as a biomarker. The spectra were processed to be statistically comparable. This option was implemented using the MetaboAnalystR package running in RStudio MetaboAnalystR package running in RStudio (27). Hierarchical clustering was used for the inspection of biochemical blood tests. In hierarchical cluster analysis (HCA), samples are combined until each belongs to a cluster measured by similarity. In this instance, we used Euclidean distance as the similarity measure and Pearson's correlation as the rank correlation. The clustering algorithm was average linkage. In addition to dendrograms, a heat map was used as a visual representation. Hierarchical clustering was performed with the*hclust* function in package *stat* in RStudio (27).

RESULTS

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

Analysis of all study groups

Figure 1 A and B shows that the PLS-DA model based on the four group was able to distinguish the two major groups (fetum and mother plasma), but does not allow a proper group separation among the pre-eclampsia and control groups related to each of the major groups.

Considering that the 5 components of the PLS-DA model (Figures 1 A and B) led to high values of R^2 (0.83), Q^2 (0.77) and a low value of accuracy (0.55), our results suggest that the PLS-DA model is able to predict whether a sample belongs to a determined group (fetum or mother plasma), but cannot which distinguish if the plasma belongs to the control or pre-eclampsia group. Thereby, further analysis should be performed in these major groups to allow the use of FTIR as an effective predictive diagnostic tool for pre-eclampsia diagnosis.

Based on the FTIR band assignment and VIP heat maps, the fetum plasma shows higher concentration of proteins, phosphorylated molecules (e.g., DNA and phospholipids), but lower concentration of carbohydrates (peak 3) and lipids (peak 2) when compared with the pregnant groups (GControl and GPre) (Figure 2). Notwithstanding, the bands assigned as 3 and 2 in the FTIR spectra are clustered in the same node in the dendrogram in the heatmap, emphasizing how these bands supports the predictive model to separate those samples from the two different major groups (fetum and mother plasma).

Analysis of fetuses versus pregnant groups

When analyzing data of all 4 study groups (FControl, FPre, GControl and GPre) at the same time, our predictive model favored the discrimination between fetum and mother plasma groups. If separating the 4 groups into fetuses (FControl + FPre) and pregnant (GControl + GPre), the two new groups can be reliably classified by the PLS-DA model with 99.7 % accuracy. This accuracy can be achieved by using only 10 wavenumbers (Figure 3).

Analysis of fetus control and pre-eclampsia groups

When the healthy fetal plasma is compared with that of a pre-eclampsia fetal plasma, it was not possible to build a predictive model able to discriminate both groups with accuracy higher than 63.3 % by using PLS-DA. Figure 8S indicates that it is not possible to significantly increase the predictive accuracy of the model, not even by increasing the number of features (wavenumbers) in the model. Thereby, the sensitivity and specificity of the model (graph on the right) were not above 74% unless at least 100 features are considered. It is worth noting that the small number of samples in each group (14 in FPre and 13 in Fcontrol) unfavor the development of a more accurate predictive model until more samples are included in the statistical analysis. Once more samples are included, higher classification performance metrics might be achieved (Figure 4).

Figure 9S shows that a separation between FControl and FPre starts to be clear by using PC1 (component 1) and PC2 (component 2) and becomes even clearer when including PC3 (component 3) despite a few outliers from the group FPre being found at threshold values of the Fcontrol group. Despite the small specificity and sensibility values achieved, Figure 10S shows that the model exhibits satisfactory performance in terms of the variables accuracy and R^2, which correspond to the ability of building a model as from real data. Although Q^2 performance was low, values around 0.4 are reasonable for biological samples. Also, Q^2 is related to the predictability of the model.

According to the heatmap of Figure 11S of the supplementary material, the main difference between the FControl and FPre groups can be observed between 900 and 1200 cm-1, which includes FTIR spectral region of phosphorylated and carbohydrates species. Based on analysis of single wavenumbers best for sample classification (Figure 4), we found that the main FT-IR spectral differences according to PLS-DA occur at wavenumbers corresponding to the absorbance of carotenoids, DNA, RNA, proteins (including collagen), lipids and fatty acids (Table 3).

Several physiological mechanisms may be associated with the biochemical changes of plasma samples. Epigenetic changes may be related to PE, as methylation of DNA, aberrant miRNA expression and histone

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

modification are the most relevant aspects related to epigenetic changes (29-31).

Altogether, these epigenetic changes could be associated with an increase of cell-free DNA and RNA in the fetus plasma. In addition, an increased volume of DNA and RNA may arise from the placenta to the mother's plasma has been reported, which explains the higher levels of phosphorylated molecules in the FPre groups (29-31). Carotenoids can block oxidative stress through cellular signaling, activating enzymes and antioxidants. The concentration of carotenoids in woman plasma and placenta is lower in women with pre-eclampsia, which agrees with our data. Besides, lipids and fatty acids levels are lower in woman with PE, characterizing a disorder in lipid metabolism (32,33) (Figures 5 e 6).

Analysis of pregnant control and pre-eclampsia groups

The discrimination between the pregnant control (Gcontrol) and pre-eclampsia (GPre) groups based on the mother's plasma was less clear than the discrimination between fetal plasma samples (Figures 7, 12S and 13S). Our results suggest that the changes in the mother's plasma are not significant enough to be sensitive in the FTIR analysis.

Plasma volume is reduced among women with pre-eclampsia. Thereby, it is reasonable to assume that their plasma is more concentrated, causing an increase in absorbance in their plasma spectra when compared with the control group. Some individuals from the control or the pre-eclampsia groups did not fit within their groups, and further analysis is needed to understand whether the statistical model has not been sufficiently optimized, or they have specific features (e.g., time of gestation).

Some circulating fetal RNA in maternal plasma can be up to 10-fold more abundant in pre-eclampsia women than in the control group, such as the mRNA coding for corticotrophin-releasing-hormone (CRH) (34,35). In fact, a recent study (2022) showed that RNA profile in woman plasma can be used to diagnose PE with 75 % sensitivity and a predictive value of 32.3 % (34). In addition, increased protein and DNA or RNA has been observed in woman with PE (34-36).

Lipid metabolism/transport proteins and ECM proteins have been upregulated in early-onset and late-onset PE, respectively (37-39). Besides, peptides expression profiles in the placenta have shown distinct patterns in normal and PE pregnancies (38). Moreover, circulating microparticles proteins have also been used to predict the risk of preeclampsia. Altogether, these results suggest that changes in protein levels in blood plasma can display potential use as biomarkers (39). Our work, however, shows that peptide and protein concentration in plasma is higher in PE than in control groups (Table 3).

Discussion

Principal comments

The introduction of a new ancillary method for diagnosing pre-eclampsia using FTIR spectroscopy has the potential to offer a non-invasive and sensitive approach to diagnosing the condition. However, it is essential to carry out further studies to validate the effectiveness of the method, considering the ethical implications, and mainly educating healthcare professionals and considering the cost-benefit in widespread clinical implementation. FT-IR spectroscopy is a highly scientific technique for analyzing organic and inorganic materials based on the absorption of infrared radiation. Biofluid samples do not require prior preparation, the analysis time is considered fast compared to some diagnostic methods, facilitating analysis.

Clinical implications

Pre-eclampsia is a major cause of maternal and perinatal mortality and morbidity, particularly in low- and middle-income countries. Pre-eclampsia is a multisystemic disorder of pregnancy. It is characterized by varying degrees of placental maperfusion with release of soluble factors into the circulation. The clinical presentation is highly variable, but hypertension and proteinuria are commonly seen. Placental disease can cause fetal growth restriction and stillbirth. Therefore, although the exact understanding of the mechanisms underlying pre-eclampsia remains incomplete, some scientific studies have investigated the biochemical and

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

molecular changes associated with this condition, including aspects related to DNA, RNA, proteins (including collagen), and lipids.

Fibrosis is a significant histologic alteration that occurs in the preeclamptic placenta and may depend on excessive deposition of collagen I (40). In another investigation, the authors reported increased levels of fetal DNA and RNA originating from the placenta, one of the most affected organs in pregnancies complicated by pre-eclampsia (41). We found that the principal FT-IR spectral differences according to PLS-DA occur at wave numbers corresponding to the absorption of carotenoids, DNA, RNA, proteins (including collagen), lipids and fatty acids.

The feasibility of PE detection was evaluated by classification accuracy. It was demonstrated that the blockade of aquaporin-9 is associated with the expression of pro-apoptotic cells. Thus, AQP9 is associated with lactate that is transported to the placenta, leading to oxidative stress, so it may play an important role as a scavenger of free radicals in energy metabolism. The function of AQP9 is still unknown, but after its blockade and the difficulty of lactic acid entry, it caused an increase in the pro-apoptotic protein in placentas with PE, leading to cell death. These biochemical processes are directly related to mitochondria, since this organelle makes most of the oxidative reactions of cells for the generation of ATP, the main cellular component for oxygen-dependent reactions is the entry of lactic acid into cells, thus transported by AQP9, proving its importance so that oxidative stress does not occur to the mitochondrial content. Important markers such as reduced GLUT1 expression and decreased process such as glycolysis were observed in preeclamptic placentas, thus it is observed that oxidative stress can generate apoptosis and worsen the condition and promote the accumulation of reactive oxygen species and generate free radicals increasing the spillage of apoptotic aggregates into the maternal circulation resulting in systemic endothelial dysfunction (42,43).

Once diagnosed, the aim of treatment is to prevent maternal-fetal complications such as placental abruption, stroke, acute pulmonary edema, renal failure and, in severe pre-eclampsia and eclampsia, fetal deterioration, preterm delivery and neonatal respiratory distress. Based on recent literature reviews, none of the available clinical tests have reached an ideal sensitivity (>90%) for predicting PE. Only Doppler, performed between 20-24 weeks, showed a sensitivity of $> 60\%$ for the detection of PE, especially when performed in pregnant women at increased risk in the second trimester, and for the prediction of early-onset PEG (12,44). The unification of sample processing methodology and data processing technology will be beneficial and may lead to radical improvements in their clinical applications in the form of next-generation diagnostic tools (45).

The evaluation of biomarkers for PE has been the subject of numerous studies and may be useful in the early diagnosis of PE. Ideally, the biomarker should be easy to perform and cost effective, in addition to allowing the detection of the specific hypertensive disorder of pregnancy (DHEG) as early as possible, preferably in the first trimester, even before the onset of hypertension. Recognize local specificities and adopt interventions based on the best scientific evidence available to improve prevention strategies, early detection of the condition and reduction of maternal and perinatal harm (15).

Research implications.

Thus, the differentiation of spectra can provide new information for a possible new screening protocol for risk assessment by biomarkers. Countries with lower rates of development have a greater chance of the population developing risks in the face of pregnancy, in Brazil about 23% of maternal deaths are caused by hypertensive conditions related to arterial hypertension. Therefore, the evaluation of biomarkers that may predispose to pre-eclampsia is extremely important for health professionals to prevent and thus reduce the number of conditions that develop to the death of pregnant women (15).

The assessment of biomarkers for Pre-eclampsia (PE) has been the subject of numerous studies and can be valuable in the early diagnosis of PE. Ideally, the biomarker should be easy to perform and cost-effective, in addition to enabling the detection of specific hypertensive disease of pregnancy (DHEG) as early as possible, preferably in the 1st trimester of pregnancy, even before the onset of hypertension. Recognizing local specificities and adopting interventions based on the best available scientific evidence to enhance prevention

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

strategies are crucial. This approach not only improves prevention and early detection strategies for the disease but also plays a fundamental role in reducing maternal and perinatal damages.

Limitations

The limitations of the technique are limited to what can be observed in the quantification of molecules, where other analysis techniques can be used, and the cost of the equipment. We also emphasize that more studies using the technology in a clinical setting and with larger numbers of samples are needed to consolidate the technique and the knowledge of the medical team, which will be of real benefit to patients.

Conclusion

In our study, we demonstrated that FTIR is a potential tool for the diagnosis of pre-eclampsia. The statistical model derived from PLS-DA shows that maternal and fetal plasma have different spectra. However, the fetal plasma seems more promising for the development of a classification model to diagnose pre-eclampsia, considering that higher accuracy, predictability (Q2) and correlation between samples and model (R2) were obtained using the fetal plasma. Nevertheless, if more fetal plasma from the control and pre-eclampsia groups are added to the model, a higher predictability index (Q2) can be obtained to create a more accurate model.

Optical spectroscopy is a versatile, rapid and non-destructive analytical technique for the analysis of biological tissues and biofluids. Fourier transform infrared (FT-IR) spectroscopy is based on the absorption of radiation by sample molecules and provides information about their structure and composition. Most of the molecular-specific information is concentrated in the fingerprint region between 600 and 1800 cm- 1 investigated in this study, which could provide valuable information to clinical/medical professionals via optical biopsy (i.e., diagnosis of lesions and pathologies characterized by molecular-specific biomarkers present in human tissues, cells, and secretions).

Author Contributions: S.M.S.D.S, S.L.S and S.C.C. were involved in investigation, data acquisition, and writing original draft., M.S.N was involved in methodology, formal analysis, investigation, interpretation, validation, visualization, writing – original draft, writing – review & editing, supervision, resources, project administration, R.B. and H.S.M. was involved in software development, formal analysis, visualization, resources, and writing original draft, R.A.S. was involved in writing – original draft, writing – review & editing, supervision, resources, project administration, and L.F.C.S.C. was involved in conceptualization, experimental design, formal analysis, investigation, software development, Validation, Visualization, Writing – original draft, Writing – review & editing, supervision, resources, project administration, funding acquisition.

REFERENCES

TABLES

Table 1. Patients demographics

Variable	n or mean	SD
Total patients	30	
Age (y)	27	5.44
Gestacional age at delivery (wk)	39.42	1.23
Birthweight (kg)	3.26	0.044
BMI ate delivery (kg/m^2)	29.3	8.1
30	12	
>30	18	
Race	Race	Race
White	22	
African American	6	
Hispanic	3	

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

Variable	n or mean	SD
Parity	Parity	Parity
Nulliparous	11	
Multiparous	19	
Labor	Labor	Labor
Spontaneous	4	
Induction	26	
Outcome		
Veginal delivery	23	
Casarean delivery	7	
Induction or augmetation methods	Induction or augmetation methods	Induction or augmetation methods
Pitocin	26	
Misoprostol	15	
Foley bulb	10	
Amniotomy	17	
Comorbidities	Comorbidities	Comorbidities
Preeclampsia	2	
Gestacional diabetes mellitus	4	
Gestacional hypertension	1	
Chronic hypertension	4	

Table 2. Biochemical components allowing for the most accurate classification of the FControl and FPre groups, and corresponding to peak wavenumbers of FT-IR bands to which these components were assigned.

Peak (cm-1)	Assignment
950	Carotenoids
971	vPO2- from proteins, DNA, RNA
1083	Symmetric PO2-
1186	Collagen and proteins
1769	v(C=O) lipids, fatty acids

Table 3. Biochemical components allowing for the most accurate classification of the 4 study groups and corresponding to peak wavenumbers of FT-IR bands to which these components were assigned.

Peak (cm^{-1})	Assignment
1186	Amide III and deoxyribose
1298	Deformation of N-H bonds in proteins
1205	νPO2- bonds, C-O bonds in polysaccharides or Amide III
1329	δ(CH), ring (polysaccharides), streching C-N in proteins

Table 4. Biochemical components allowing for the most accurate classification of the GControl and GPre groups and corresponding to peak wavenumbers of FT-IR bands to which these components were assigned.

Peak (cm1)	Assignment
981	vPO2- from proteins, DNA, RNA
1083	Symmetric PO2-

Peak (cm1)	Assignment
1104	Symmetric streching of P-O-C bonds
1544	Amide II bands
1657	Amide I (vC=O, δC-N, δN-H)

FIGURES

Figure 1. A) PLS-DA scores for the 4 study groups: control pregnant women (GControl; dark blue) and pre-eclampsia pregnant women (GPre; light blue), control fetuses (FControl; red), pre-eclampsia fetuses (FPre; green). B) Performance of each of the 5 principal components (PC) of PLS-DA according to accuracy, Rˆ2 and Qˆ2 values.

Figure 2. Heatmap of variance importance projection scores varying over a color scale from dark blue (-4) to dark red (+4). Groups were separated according to the dendrograms on the top and the lateral (left side) of the heatmap. The left side of the top dendrogram represents the separation between control pregnant women (GControl; dark blue) and pre-eclampsia pregnant women (GPre; light blue), whereas the right side represents the separation between control fetuses (FControl; red), pre-eclampsia fetuses (FPre; green). Similarly, each portion of the lateral dendrogram shows the separation of groups of wavenumbers most useful to classify samples. The right side of the heatmap shows the wavenumbers grouped by colors corresponding to spectral regions of vibrational modes corresponding to lipids (blue), proteins or peptides (green), phosphate functional groups (pink), and carbohydrates (yellow).

Figure 3. Mean blood plasma spectra for the GControl (blue line) and GPre (green line), FControl (black line), FPre (red line). Biochemical components were assigned to spectral regions of their corresponding vibrational modes. This assignment follows the same color code of the heatmap.

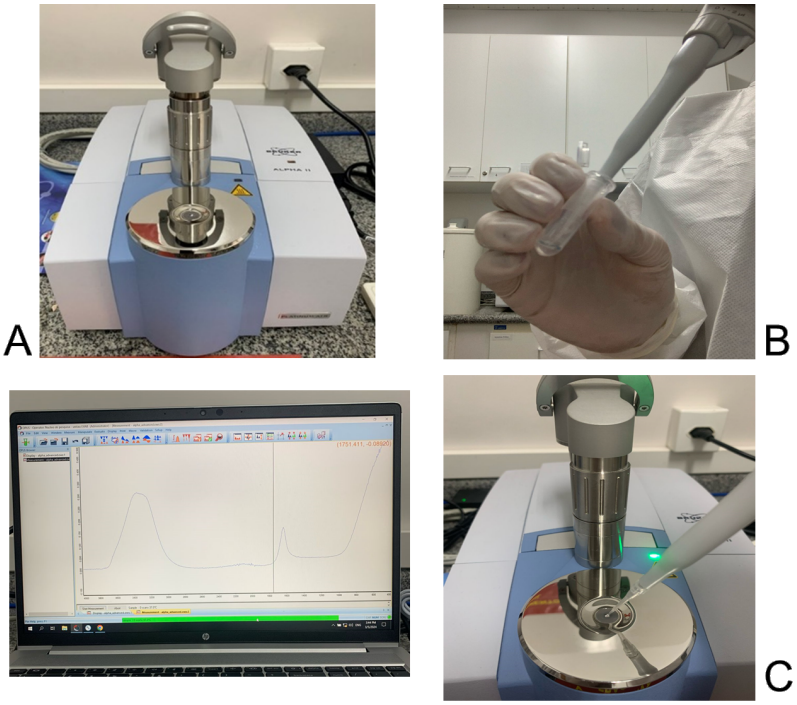
Figure 4. A) PC1 and PC2 scores for the control fetuses (FControl; red), pre-eclampsia fetuses (FPre; green). We could already observe only a good separation between study groups by using only the first two PCs of PLS-DA. B) PC1, PC2, and PC3 scores for the FControl and FPre groups. The study group separation was even better by considering 3 PCs.

Figure 5. Receiver operating characteristic (ROC) curve of the PC scores of PLS-DA for each of the 5 wavenumbers leading to highest classification performance. The boxplots on the right side of every ROC curve represent the PC scores for FControl and FPre groups. The red line shows the threshold for the best separation between the two groups. ROC curves are shown for PC scores of the following wavenumbers: A) 971 cm^{-1}, B) 1084 cm^{-1}, C) 951 cm^{-1}, D) 1770 cm^{-1}, and E) 1186 cm^{-1}.

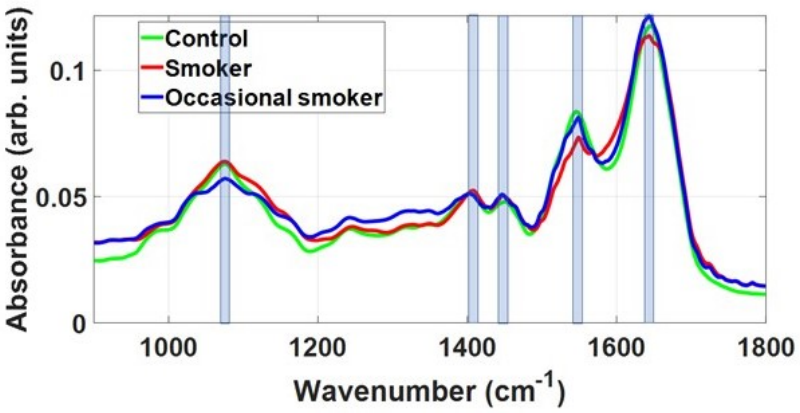
Figure 6. A) Receiver operating characteristic (ROC) curve of the PC scores of PLS-DA for each of the 5 wavenumbers leading to highest classification performance. The boxplots on the right side of every ROC curve represent the PC scores for GControl and GPre groups. The red line shows the threshold for the best separation between the two groups. ROC curves are shown for PC scores of the following wavenumbers: A) 1084 cm^{-1}, B) 981 cm^{-1}, C) 1104 cm^{-1}, D) 1657 cm^{-1}, and E) 1545 cm^{-1}.

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57033838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.



Posted on 5 May 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.171492455.57633838/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.



Bands (cm^{-1})	Vibrational modes	Structural (Biomolecular) components
1076	Symmetric phosphate [PO_2] stretching	DNA
1403	CH_3 bending of methyl groups (proteins) and $\delta_s CH_3$ (collagen)	Proteins and collagen
1451	Asymmetric CH_3 bending modes of methyl groups	Proteins
1547	Amide II	Peptides and proteins (Amide II)
1646	Amide I, C==O (C_5 methylated cytosine), stretching C==C (uracil), NH_2 guanine	Peptides and proteins (Amide I), methylated cytosine C_5, uracil, and guanine