A metabolomics-based approach to the screening of endometrial cancer: development of a gas chromatography-ion trap mass spectrometry-based method

Allycia Lee

University of Tennessee at Chattanooga, allycialee@gmail.com

Follow this and additional works at: https://scholar.utc.edu/honors-theses

Part of the Pathogenic Microbiology Commons

Recommended Citation

A metabolomics-based approach to the screening of endometrial cancer: development of a gas chromatography-ion trap mass spectrometry-based method

Allycia Yuen Kay Lee

Departmental Honors Thesis
The University of Tennessee at Chattanooga
Department of Chemistry & Physics

Examination Date: April 8, 2021

Steven Symes
UC Foundation Professor of Chemistry
Thesis Director

Wang-Yong Yang
Assistant Professor of Chemistry
Department Examiner
# Table of Contents

1  Abstract .......................................................................................................................... 2

2  Introduction .................................................................................................................... 3

3  Methodology ................................................................................................................... 5
   3.1  Extraction .................................................................................................................. 5
   3.2  Purification and Solvent Removal ............................................................................... 7
   3.3  Derivatization .......................................................................................................... 8
   3.4  GC-MS Analysis ...................................................................................................... 10

4  Method Optimization ....................................................................................................... 13
   4.1  Standard Analysis .................................................................................................... 13
   4.2  Whole Blood and Serum Analyses .......................................................................... 19

5  Discussion ...................................................................................................................... 22
   5.1  Sample Preparation ................................................................................................. 22
   5.2  Instrumental Method ............................................................................................... 24
   5.3  Clinical Trial Potential ............................................................................................ 26
   5.4  Metabolomics Standards Initiative .......................................................................... 26
   5.5  Future Work ............................................................................................................ 27

6  Conclusion ..................................................................................................................... 28

7  References ..................................................................................................................... 29
Abstract

Endometrial Cancer (EC) is the most common malignant tumor of the female reproductive tract. Currently there are no methods for effectively screening EC, but one method that is thought to be useful in screening EC is metabolomics. Metabolomics involves the analysis of metabolites which are low molecular weight inorganic and organic chemicals that are substrates, intermediates, and by-products of enzyme-mediated biochemical reactions in the cell. Previous research shows that the metabolic signature of EC patients are discernable from those of healthy patients. To aid in the development of an effective screening method for EC, blood serum, urine, and saliva samples of EC patients and controls enrolled in a clinical trial at the University of Tennessee College of Medicine at Chattanooga Department of Gynecological Oncology will be collected. For each sample, the metabolites will be extracted, purified, and derivatized before gas chromatography-mass spectrometry (GC-MS) analysis. Chromatographic separation will use a Trace1310 GC system equipped with a TG-5MS column interfaced with ion trap mass spectrometry. The obtained signatures will be used in an effort to build multivariate classification models that can distinguish EC patients from healthy patients, providing a non-invasive, affordable screening test for EC. Prior to the analysis of clinical samples, the entire method will be optimized for all sample types to ensure that the method will be able to provide the data necessary for developing a screening system. The initial phases of method development indicate that the proposed method will be sufficient for collecting metabolomic data, but more optimization is required before clinical samples may be analyzed.
Introduction

Endometrial cancer (EC) is the most common malignant tumor in the female reproductive tract and the sixth most common neoplasm in women worldwide.\(^1\) Over the past ten years, the severity and incidence of EC has increased at an average rate of 1-5% each year.\(^1\) This increase is associated with the rising incidence of certain risk factors for EC, including obesity and type 2 diabetes. Currently, there are no methods for effectively screening EC, so most patients are unaware of the cancer until they develop cancer-related symptoms or have had invasive surgery in the genital tract. Unfortunately, even the detection of EC via invasive procedures is not completely accurate.\(^2,3\) Because of the difficulty of detection, a patient’s 5-year survival rate for EC would likely have decreased significantly by the time the cancer has been noticed as the cancer may have progressed past the initial stages.\(^4\)

One method that is thought to be useful in screening EC is metabolomics. Metabolomics is the qualitative and quantitative analysis of metabolites which are low molecular weight (< 1.5 kDa) inorganic and organic molecules that are substrates, intermediates, and by-products of enzyme-mediated biochemical reactions in cells.\(^5\) Metabolites are involved in essential cellular functions and can be produced by the individual’s metabolism or be derived from normal flora, diet, and other external sources.\(^6\) Cells, tissues, serum, urine, and most other biological fluids may all be utilized in the detection of metabolites.\(^7\) The biological significance of metabolomics lies in the fact that metabolites represent the actual state of the body in comparison to other “-omics” sciences such as genomics and proteomics. While genes and proteins may or may not be expressed or translated, respectively, under a given set of conditions, the presence and quantity of a particular metabolite necessarily implies whether or not an upstream process occurred, and thus their pattern is a direct reflection of the observed phenotype.\(^8\) Metabolomics is
better for pattern recognition in biological samples than any other post-genomic technology.\textsuperscript{9} Many cancers and other diseases have been studied using metabolomics to better understand the disease, aid in the detection of the disease, and/or aid in the development of treatment options.\textsuperscript{10-16}

The development of an effective screening system for EC using metabolomics will provide a non-invasive, affordable screening test for EC. The basis behind a metabolomic-based screening system is biomarker discovery. By identifying specific metabolites that are impacted by a disease state, metabolomics can be utilized to quickly identify whether or not an individual is being affected by a given disease or if a specific medication is working. A prior study conducted by Troisi et al. (2018) has shown that there are differences in the metabolic signature of EC patients that are discernable from those of healthy patients and could also be used to distinguish between the two histotypes of EC.\textsuperscript{17} Lactic acid, progesterone, homocysteine, 3-hydroxybutyrate, linoleic acid, stearic acid, myristic acid, threonine, and valine were the metabolites that were determined to be the most significant in distinguishing EC patients from controls.\textsuperscript{17} This previous data, however, does not take into account possible differences in the metabolome from individual characteristics and was only performed on an Italian cohort. As such, further studies are necessary to confirm the previous findings. By matching cases and controls with respect to age, sex, ethnicity, genetic background, and socio-economic conditions, the possible effects of confounding factors will be minimized, resulting in a more effective distinction between the diseased and non-diseased state. Accounting for these factors is crucial as these factors can affect an individual’s metabolic signature.\textsuperscript{18,19}

Gas chromatography-mass spectrometry (GC-MS) is an effective method for analyzing the metabolic signatures of biological samples. A clinical trial is being conducted to determine
the metabolic signatures of healthy post-menopausal subjects and post-menopausal EC patients enrolled with the University of Tennessee College of Medicine at Chattanooga Department of Gynecological Oncology located at Erlanger Hospital. We will use GC-MS to measure the metabolic signatures of the samples collected from the women enrolled in the clinical trial. Serum, urine, and saliva samples will be collected for each enrolled subject. The obtained signatures will be used to build multivariate classification models that may be able to distinguish EC patients from healthy patients. Before the analysis of clinical samples can occur, the method for preparing and analyzing the samples must be optimized for the instrumentation that will be utilized. Such method development is crucial in ensuring the reliability of the data to be collected. The current study addresses the initial phases in method development and the overall potential for the proposed clinical trial.

**Methodology**

The methodology in which the collected samples would be analyzed under is a very important factor in the success of the study as the method will dictate the quantity and type of metabolites that could be detected. The sample preparation method is based on the method from the MetaboPrep GC kits by Theoreo, and the overall methodology is similar to methods utilized by Troisi et al. (2018) and in other GC-MS-based metabolomic experiments.\textsuperscript{15-17,20-21} Each step of sample preparation and the conditions of the GC-MS method presented in this thesis are the most updated method resulting from our first attempts at method optimization.

**Extraction**

For analysis of any sample type that contains cells, metabolite extraction is the first step in sample preparation. This process aims to release metabolites from within a sample’s cells and
increase the quantitative yield of metabolites. The extraction method is important to optimize for a given study because different approaches can lead to contradictory interpretation of the resulting metabolomes. The primary method of extraction is liquid-liquid extraction (LLE) which utilizes organic solvents to precipitate higher molecular mass biochemicals, such as proteins, RNA, and DNA. The samples are centrifuged, and the precipitated components are discarded while most, if not all, metabolites will remain in the supernatant solution. It should also be noted that two extractions can help to increase the metabolite yield, but more than two extractions is likely to decrease metabolite yield.

The specific type of LLE that will be utilized for the current clinical trial is a monophasic extraction, which involves one miscible solvent system. These types of extraction solutions are well suited to extract a wide range of metabolites. Monophasic extraction is often used for cell cultures, tissues, and some biofluids. Typical combinations of organic solvents used are methanol/water, water/acetonitrile, and methanol/water/chloroform. The exact solution to be utilized in this work is a mixture of methanol, deionized water, and chloroform in a 2.5:1:1 ratio. The extraction solution also contains 2-isopropylmalic acid at a concentration of about 111ppm (25mg/225mL) to serve as an internal standard.

For the extraction process, 50 μL of the biofluid sample is added to 200 μL of the extraction solution in a 2 mL Eppendorf tube. The mixture is then mixed for 30 minutes at room temperature on a tube revolver. After 30 minutes, the sample is centrifuged for 10 minutes at 12500 rpm at 4℃. 200 μL of the supernatant is then transferred to another 2 mL Eppendorf tube that contains the purification solution.
**Purification and Solvent Removal**

Following the addition of the 200 μL of the supernatant to the purification solution, the sample is vortexed for 30 seconds and then centrifuged for 10 minutes at 12500 rpm at 4°C. 300 μL of the resulting supernatant is transferred to a clear 2 mL screw cap glass vial suitable for analysis by autosampler.

The purification solution (PS) utilized in the sample preparation is a 10 mM solution of tetrabutylammonium bisulfate with the same concentration of the internal standard as the extraction solution. The purification solution contains internal standard because some biological samples, such as cerebrospinal fluid or amniotic fluid, do not contain cells. These samples would be added directly into the purification solution without extraction. The tetrabutylammonium sulfate serves as an ion pairing agent. Ion pairing agents are compounds that possess an opposite charge to that of the charged or polar metabolites of interest in an aqueous solution. They serve as a means of improving the yield of polar metabolites and are more frequently used in LC-MS experiments. Tetrabutylammonium is an ion pairing agent that is beneficial for the analysis of anionic metabolites.

Immediately following purification is solvent removal. At this stage, the solvent is composed of a mix of organic solvents with water. The solvent must be removed as the water or organic solvents used could cause problems in later steps or during analysis. In this case, the removal of the solvent also serves as a sample concentration step. The different approaches for drying a sample include the use of a lyophilizer or a N\(_2\) evaporator. Both methods sufficiently remove the solvent. For lyophilization, the sample is frozen at -80°C before being lyophilized overnight. The lyophilizer used is a Buchi Lyovapor L-200. For N\(_2\) evaporation, the samples are placed on a rack under steady N\(_2\) flow until the solvent has completely evaporated. This can be
completed in a couple of hours at most. Solvent removal is usually performed for GC-MS analysis because as much water as possible needs to be removed from the samples so as to not interfere with the derivatization of the metabolites.

*Derivatization*

The last step of sample preparation is derivatization. Derivatization is perhaps the most important aspect of sample preparation for GC-MS analysis since most metabolites are polar and heat-labile. The process of derivatization replaces polar (active) hydrogens on metabolites with a non-polar substituent to prevent H-bonding and increase the volatility of the metabolites as a result. The functional groups that are most commonly affected by derivatization are –OH, –COOH, –SH, –NH₂, and –CONH. Derivatization of metabolites needs to be optimized in individual laboratories as the derivatization conditions can impact the limit of detection (LOD), sensitivity, and selectivity of the GC-MS. This process can also improve chromatographic separation and stability of the metabolites. The derivatization occurs as a two-step process.

The first step of derivatization is methoximation which involves a reaction with methoxyamine hydrochloride, which is typically dissolved in the aprotic solvent pyridine at a concentration of 20 mg mL⁻¹. 50 μL of the methoxyamine solution is added to the glass vial containing the dried metabolites and is mixed for 90 minutes on a tube revolver. During this time, the methoxyamine reacts with ketones and α-ketoacids to serve as a protecting agent and prevent decarboxylation (*Fig. 1*). The methoxyamine may also react with sugar tautomers to cause ring opening and protects aldehydes and ketones (*Fig. 2a*). This reaction, paired with the second step of derivatization, reduces the number of sugar tautomers that form during derivatization, lowering the number of chromatographic peaks which facilitates better separation and quantification of metabolites.
Figure 1. The methoximation of propanone protects the carbonyl from further reactions.

The second part of derivatization is trimethylsilylation derivatization (Fig. 2b). In this process, active hydrogens are replaced by trimethylsilyl (TMS) groups to produce TMS derivatives. The two most commonly used TMS reagents are $N, O$-bis(trimethylsilyl)-trifluoroacetamide (BSTFA) and $N$-methyl-$N$-trimethylsilyltrifluoroacetamide (MSTFA). BSTFA is the derivatization reagent for trimethylsilylation used in this work and is generally a better silyl donor than MSTFA. Once the 90 minutes of methoximation had elapsed, 25 μL of BSTFA in 1% trimethylchlorosilane (TMCS) is added to the glass vial and mixed for another 90 minutes. The TMCS in this solution serves as a catalyst for the reaction.

Dryness, as well as solvent and sample purity, have the largest impact on the success of the derivatization, whereas derivatization time and temperature can vary widely and should be optimized in each laboratory. In terms of dryness, the amount of water in the sample must be as minimal as possible prior to derivatization. Water interferes with the silylation process because water will react more quickly with the derivatization reagent than the metabolites. The derivatization of water forms trimethylsilanol (Fig. 2c), which separates into a separate solvent layer, impeding the derivatization of metabolites. This becomes important when considering that metabolite extraction and purification steps use aqueous solutions, and water constitutes the majority of all biofluid sample types. Derivatization is severely hampered in samples with a water content of more than 10%. As a result, the sample dry-down step is completed very carefully prior to addition of derivatization reagents.
Figure 2. The (a) methoximation and (b) BSTFA trimethylsilylation of glucose opens the cyclic form and removes the active hydrogens. (c) The trimethylsilylation of water prevents the derivatization of metabolites.

At the end of derivatization, the 75 μL samples are transferred into a glass insert to facilitate autosampler injection since 75 μL in the bottom of a 2 mL vial will not be reachable by the autosampler syringe needle. The insert is centrifuged for 10 minutes at 12500 rpm at 4°C and then placed back in its respective 2 mL vial and stored at -20°C or below until analysis.

GC-MS Analysis

The most prominent methods for analyzing metabolites are nuclear magnetic resonance and mass spectrometry (MS) paired with chromatography. Among these, MS is becoming the preferred analytical method because it offers the most sensitive detection and broadest
metabolome coverage, being able to detect hundreds of metabolites with a single method.\textsuperscript{23,31} Though MS does provide the best metabolite coverage, MS-based methods will only be able to detect metabolites that can be ionized. MS is usually directly preceded by separation techniques with gas chromatography (GC) and liquid chromatography (LC) being the most frequently used.

There are a variety of mass spectrometers that can be used for metabolomics. The best mass spectrometer for a given experiment will depend on whether an experiment will have a targeted or untargeted approach. For an approach targeting specific metabolites, a triple quadrupole is preferred because you can monitor for the appearance of parent/product fragmentation reactions, also called selected reaction monitoring (SIM), to aid in the identification of specific metabolites.\textsuperscript{23} For untargeted approaches, like the current clinical trial, the high sensitivity of a single quadrupole mass spectrometer makes for a highly capable instrument, despite the lower resolution compared to some other mass spectrometers, such as time-of-flight (TOF) or Orbitrap mass analyzers. Both TOF and Orbitrap have substantially higher mass resolution with the ability to determine “exact” masses and are powerful platforms that are used for targeted and untargeted experiments.\textsuperscript{23,32} Tandem MS/MS systems offer the highest possible selectivity and are best suited for the identification of unknowns. Generally, any of these mass analyzer systems would allow for specific, sensitive, and quantitative metabolite analysis. For GC specifically, a single quadrupole is very effective, but a TOF mass analyzer could also be used if higher mass resolution is needed.\textsuperscript{33} In the current study, the mass analyzer to be used is a quadrupole ion-trap (QIT). In comparison to the other mass analyzers mentioned above, QIT has considerably lower sensitivity, so the usage of this particular mass analyzer may result in a lower number of detected metabolites than other mass analyzers. Fortunately, QIT mass analyzers still have the potential for high sensitivity under proper care and offer a large
mass range with variable mass resolution. There have been several metabolomic studies that have successfully utilized QIT, but most of the time it has been used for high-throughput direct-injection or in tandem with another mass analyzer. It is most frequently combined with LC and electrospray ionization. Generally, QIT is not a preferred mass analyzer for metabolomics.

Gas chromatography-mass spectrometry (GC-MS) is a technique that is capable of measuring a wide range of water-soluble metabolites. In fact, GC-MS is the best analytical method for low molecular weight and volatile metabolites, often providing higher sensitivity than NMR or LC-MS detection. It can detect small and uncharged species, like very short fatty acids, sugars, acids, and alcohols, that may not be detected using LC-MS. Due to the very long column lengths, GC has the highest separation efficiency and cleanest chromatograms, which takes some pressure off of the performance of the MS. Based on the principle of how a GC operates, it is only effective for compounds that will volatilize (without decomposing) within the operating temperature. It is for this reason that the derivatization of metabolites is necessary in this method to maximize metabolite coverage. As many of the compounds within the sample must be volatile because non-volatile compounds can degrade inside the GC inlet and cause cross-contamination of samples. Hard ionization techniques, such as electron impact (EI), are often used with GC-MS to create many reproducible fragments that aid in identification of metabolites but will produce few, or no, molecular ion peaks.

The GC-MS system for the clinical trial is a Trace 1310 GC interfaced with an ITQ 1100 MS and an AS 1310 autosampler. The GC column that is utilized is a ThermoFisher TG-5MS column (30m x 0.25mm x 0.25um). The inlet of the column is placed in a splitless liner. For analysis, the autosampler draws 1 μL of the sample and injects it into the GC with a 6 second pre-injection and post-injection hold time in the inlet. The initial temperature of the oven is
100°C (hold 1 min) and ramps up to a final temperature of 320°C (hold 2.33 minutes) at 6°C per minute, resulting in a total run time of about 40 minutes. The flow of the column is held constant at 1.1 mL/min with an inlet temperature of 280°C and a split flow of 5.5 mL/min. The temperature of the transfer line leading to the MS was also set at 280°C. For the MS, ionization of the eluents occurs by 70 eV EI. The ion source has a temperature of 225°C. Scans occur beginning at 4.0 minutes and scans from 50-505 m/z with 1 microscan. Under these conditions, the scan time is 0.21 seconds.

**Method Optimization**

Thus far, the majority of the method development has focused on the instrumental method with only some considerations for the optimization of sample preparation. Modifications of the instrumental method provides a better understanding of the instrument capabilities for a metabolomic project. As of now, there have been two distinct phases of method development. The first phase involved the testing of a standard mix of metabolites while the second phase incorporated the analysis of biological samples, namely whole blood and serum.

**Standard Analysis**

The metabolite standard that was utilized was provided by the lab that conducted the 2018 EC study along with the chromatogram obtained in their lab (*Fig. 3a*). The standard contains a total of 15 metabolites with varying molecular masses and polarities. *Table 1* provides a list of the metabolites, the ion masses used for identification of the metabolites, the concentration in the undiluted standard, and the theoretical concentration in the final sample after 1/10 dilution and sample preparation. These metabolites are presented in order of expected elution from the GC. It should be noted that the theoretical concentration will not necessarily be
reflected in the signal intensities in the chromatograms as different analytes can have dramatically different ionization efficiencies so that an equimolar solution of two metabolites could produce signals that vary by $>10^6$ ions.$^{40}$

**Table 1.** Metabolite concentration in undiluted standard and in a 1/10 diluted final sample

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Identification Ion (m/z)</th>
<th>Standard Concentration (ppm)</th>
<th>Final Concentration (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyruvate</td>
<td>174</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Succinate</td>
<td>247</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Glycine</td>
<td>248</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Lactate</td>
<td>117</td>
<td>111.6</td>
<td>4.464</td>
</tr>
<tr>
<td>Ketoglutarate</td>
<td>156</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Tartrate</td>
<td>292</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Ribitol</td>
<td>319</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Citrate</td>
<td>273</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Glucose</td>
<td>319</td>
<td>100</td>
<td>4.00</td>
</tr>
<tr>
<td>Arachidate</td>
<td>369</td>
<td>500</td>
<td>20.0</td>
</tr>
<tr>
<td>Sucrose</td>
<td>361</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Testosterone</td>
<td>389</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Colecalciferol</td>
<td>351</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Tocopherol</td>
<td>502</td>
<td>1000</td>
<td>40.00</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>458</td>
<td>1000</td>
<td>40.00</td>
</tr>
</tbody>
</table>

For the initial tests of the standard, a 1/10 dilution was prepared by mixing 45 μL of normal saline (0.9% NaCl) with 5 μL of the standard. This sample was processed following the exact instructions provided in the MetaboPrep GC kits for serum samples and was analyzed using an updated 50-minute GC-MS method that was used in the Troisi lab.$^{39}$ This 50-minute method varies in several ways from the method described previously. In this method, the final hold time was 12.33 min, and the temperature of the transfer line was 320ºC. On the MS side, the scan was from 35-600 m/z with 3 microscans with a cut time of 6 minutes and an ion source temperature of 200ºC. These tests were conducted before the addition of the autosampler to the GC-MS system, so the samples were injected by hand with a volume of 2 μL. In the first analysis
of the standard, only 10 out of the 15 metabolites could be identified in the chromatogram by spectral matching to the National Institute of Standards and Technology (NIST) Mass Spectral Library (www.nist.gov). The detected metabolites were glycine, succinate, lactate, ketoglutarate, tartrate, ribitol, citrate, glucose, sucrose, and testosterone (Fig. 3b). The same process was repeated with 3 more standard samples, and similar results were produced each time. It was

Figure 3. (a) The chromatogram from Troisi et al. (2018) detected all 15 metabolites by SIM of the identification ions from Table 1. (b) The chromatogram of the first standard sample detected 10 out of the 15.
noted that there were many peaks in the chromatogram that were unassociated to the metabolites present in the standard and were likely byproducts of sample preparation. To account for those peaks, a method blank was processed with 50 μL of normal saline for all subsequent sample analyses.

Due to several of the standard metabolites not being detected, changes were made to the instrumental method to try to increase the sensitivity and overall amount of data collected. Firstly, the cut time was lowered to 4 minutes from 6 minutes. Glycine and succinate both had a retention time of about 7.4 minutes and are the first metabolites from the list that could be identified. Theoretically, pyruvate should have eluted first, so the retention time needed to be lowered in case it eluted too early but could not be lowered too much in order to avoid overloading the detector with solvent. Next, the scan range was adjusted to 50-500m/z and the microscans were lowered to 1. Both of these changes were made to decrease the scan time from 0.56 sec to 0.21 seconds without significantly impacting the identification of metabolites by their mass spectrum. The decrease in scan time would increase the chromatographic resolution but may hamper the identification of low abundance metabolites that elute very quickly. Even with these changes, the scan time was still greater than the 0.175 second scan time of the original EC study. The mass range was later altered to 50-505m/z to account for the detection of tocopherol. The same standard samples that were prepared previously were reanalyzed under these updated conditions. More peaks did occur between 4-6 minutes, but none of the peaks could be identified as pyruvate by spectral matching to the NIST Mass Spectral Library.

At the later retention times, there was a general lack of peaks on the chromatogram after 33 minutes. The noise of the instrument was also substantially higher after the column reached its final temperature. Due to this lack of data at the higher retention times, the final hold time was
reduced to 2.33 minutes. The ion source temperature was also raised to 225°C as increased ion source temperatures are supposed to be beneficial to late-eluting analytes.\textsuperscript{41} Even with these alterations to the method, the acquired chromatograms did not vary significantly from the previous samples. As a result, the overall sensitivity of the GC-MS was questioned, leading to a full round of maintenance for the entire system to ensure that the instrument was operating as well as it could. The autosampler was added to the system around the time of the maintenance as well. The maintenance resulted in chromatograms with much higher sensitivity and a higher total ion count (TIC), allowing for cholesterol to be detected. Due to the increased sensitivity from the maintenance, the decision was made to decrease the injection volume in a repeat analysis to 1 μL to decrease the chance of overloading the detector. There was no significant decrease in the detection of metabolites from this alteration, so the 1 μL injections were continued.

Once changes to the instrumental method no longer seemed to have an effect on the chromatographic data, the focus shifted to the sample processing. To allow the standard to behave more like a serum sample, 0.100g of human albumin were dissolved in 2 mL of normal saline. The normal saline-albumin mix was used in place of just normal saline when diluting the standard. A 1/10 standard sample was prepared using the normal saline-albumin mixture and analyzed with the updated instrumental method. In the sample preparation phases, the transfer volumes were maximized in an attempt to improve the concentration of the metabolites in solution. This resulted in the transfer of 300 μL of purification supernatant to the glass vial as opposed to the 175 μL stated in the MetaboPrep instructions. This improved the TIC of the resulting chromatogram, indicating an effective increase in the concentration of the metabolites. Unfortunately, there were still standard metabolites that could not be detected. As a note, the
column condition was questioned around this point due to the analysis of a standard test mix. The column was replaced, leading to slightly longer retention times in subsequent analyses. To determine if the missing metabolites in the standard were due to low concentration or sample preparation, a $\frac{1}{5}$ dilution of the standard was also processed with maximum volume transfers while 10 μL of the standard were directly dried down and derivatized. The resulting chromatograms of both samples (Fig. 4) indicate that the sample preparation steps are not as suitable for late eluting metabolites. Arachidate and colecalciferol both appear in direct-dry standard but not the extracted sample. They both have relatively low concentration in the direct-dry sample despite having a concentration that should be over four times greater than the extracted sample. These results

![Figure 4. The chromatogram of the (a) extracted standard and (b) the direct-dry standard analyzed with the most recent method.](image-url)
imply that the concentrations of these two metabolites are too low to be detected after sample processing. Tocopherol is still undetected and could have an even lower concentration below the sensitivity of the instrument. After the changing of the column, the concentration of testosterone in the extracted sample is much lower than expected. Glycine and ketoglutarate are also not identifiable. Glycine could have coeluted with succinate. More analyses on the new column will be necessary to understand these changes.

*Whole Blood and Serum Sample Analyses*

As the analyses with the standard samples showed improvement, it was determined that a proper biological sample should be analyzed to evaluate if the current method would effectively provide metabolomic data. The first biological test sample to be analyzed with the given method was whole blood that was collected via a finger prick. Approximately 50 μL of whole blood were collected including the resulting blood clot. This sample was processed and analyzed using the method described in the methodology section. The chromatogram of the whole blood sample possessed many peaks that were indicative of unique metabolites.

After the whole blood trial proved successful, serum was the next biological sample that needed to be tested. Serum is one of the sample types of interest for the clinical trial, so this phase in the method development is crucial for measuring the potential success of the current method for the clinical trial. To collect a serum sample, the researcher had approximately 7 mL of her blood collected by a trained patient care technician. This sample was allowed to sit at room temperature for 35 minutes to clot and was then centrifuged for 10 minutes at 2860 rcf (relative centrifugal force). The serum layer (~1.750 mL) was then transferred to a pre-chilled 15 mL sterile falcon tube and placed in ice. From this vessel, the serum was aliquoted into 2 mL
glass screw cap vials in two 500 μL aliquots and one 750 μL aliquot. All three aliquots were stored at -80°C.

For the first serum trial, the 750 μL aliquot was slowly defrosted by placing the sample on ice. Once the sample was defrosted, the serum was processed in duplicate with 3 different volumes transferred to the extraction solution. The volumes used were 50 μL, 100 μL, and 200 μL with two samples processed at each initial volume resulting in 6 total samples. The volumes of the transfers were altered as a result of the increased sample volume and are listed in Table 2. The derivatization and instrumental method remained the same. The purpose of varying the starting volumes was to determine if increasing the total metabolite concentration would allow for an increase in the number of quantifiable metabolites. The duplication of each sample volume checks for reproducibility. Each of the resulting chromatograms (Fig. 5) displayed many peaks that could be associated with different metabolites. The comparison of these chromatograms of the samples does not reveal a significant difference in the number of metabolites detected. There was also a lack of consistency in the variance of the concentrations of metabolites with the exception of the increasing presence of glucose and other sugars. As a result of the lack of improvement of the chromatographic data at higher sample volumes, 50 μL will remain the sample volume. One consistent feature of all of the chromatograms was a low number of peaks above 30 minutes.

<table>
<thead>
<tr>
<th>Sample volume (μL)</th>
<th>Volume Transferred to PS (μL)</th>
<th>Volume Transferred to Glass Vial (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td>100</td>
<td>225</td>
<td>350</td>
</tr>
<tr>
<td>200</td>
<td>275</td>
<td>400</td>
</tr>
</tbody>
</table>
Figure 5. The chromatograms of one of each of the (a) 50 μL samples, (b) 100 μL samples, and (c) 200 μL samples.

There were also a few instrumental conditions that were evaluated with the serum sample. The initial method utilized a 2 μL injection of sample as did the original EC study\textsuperscript{17}, so a 2 μL injection was attempted with one of the 50 μL samples. The resulting chromatogram did not appear to have any more metabolite peaks but did double in the TIC. The issue with the 2 μL injection was that the larger peaks exhibited peak tailing and lower resolution; therefore, the 1 μL injection should be retained. The other injection condition that was altered was the pre-injection hold time. The hold time is meant to help in the volatilization of the metabolites, but it could also degrade thermolabile compounds if not careful. A decreased pre-injection hold time of 3 seconds (compared to the previous 6 seconds) did not result in changes to the chromatogram. Outside of the injection method, an increased number of microscans was tested to see if it could
improve the chromatographic data without decreasing the number of peaks. The 2 microscan chromatogram possessed smoother peaks with better peak shape without negatively altering the metabolite data. 2 microscans should be incorporated into future versions of the method.

**Discussion**

*Sample Preparation*

The method for the sample preparation has yet to be optimized at any level. Some modifications, such as the transfer volumes, have already occurred, yet the reaction time and conditions have not been evaluated and optimized. This should be the point of focus moving forward. Based on the comparison of the extracted and direct-dry standard samples, the extraction and/or purification of the samples has significant impact on the detection of metabolites. This impact is due to the fact that metabolites are extracted and/or derivatized with different efficiencies. Therefore, the conditions of these processes determine which metabolites are carried through to analysis. Optimization should aim to maximize the number of unique metabolites that could be detected using GC-MS.

Another portion of the sample preparation that could use more attention is the solvent removal. A brief test comparing standard samples prepared via lyophilization and N₂ evaporation did not indicate a significant difference in the metabolites detected. These results could use further evaluation using biological samples. It will also be important to reevaluate the results as the lyophilizer was not functioning completely as expected when it has been used previously. For the clinical trial, it would be most beneficial to optimize a method utilizing lyophilization as it would be the most effective method for processing a large number of samples simultaneously. N₂ evaporation would not be as suitable for processing more than 10-12 samples at a
time. Consistency in the dry-down method will be important, especially in terms of temperature and time.

For derivatization, the conditions can impact the LOD, sensitivity, and selectivity of the GC-MS. Some of the major concerns of the current method are sensitivity and the general lack of metabolite peaks at higher retention times, so derivatization would be the best place to start for evaluating those issues. After all, the GC-MS method optimization conducted by Danielsson et al. suggests that the optimal derivatization time is much longer than the current method. The derivatization times should be varied to determine which time range would provide a greater number of metabolite peaks without being too centered on one class of metabolites. Right now, superficial analysis of the spectral data of the first serum samples indicates that many of the peaks are likely isomers of carbohydrates. This may be a reason for concern as the numerous, large carbohydrate peaks could mask the presence of other metabolites. While methoximation is supposed to help in reducing the number of carbohydrate isomers, it may be necessary to reduce the presence of carbohydrates even more as they are very unlikely to be considered a biomarker for EC or cancers in general. One method that could be considered in the future is derivatization with \( N\text{-}tert\text{-}butyldimethylsilyl-N\text{-}methyl\text{-}trifluoroacetamide (MTBSTFA) \) rather than BSTFA. Metabolite derivatization with MTBSTFA yields \( tert\text{-}butyldimethylsilyl (TBDMS) \) derivatives, some of which may be more stable than their corresponding TMS derivatives. MTBSTFA is larger than BSTFA and MSTFA, which allows it to reduce the yield and volatility of glucose and large carbohydrates to the point that they cannot be detected. Another advantage is that TMS derivatization often yields both mono- and di-silylated amines while TBDMS derivatization is less likely to form multiple derivatives due to the bulkiness of the reagent.
**Instrumental Method**

The instrumental method has been adjusted many times, but it will still need some minor modifications as method optimization continues. Looking at the standard analyses, the method under the current state of the instrument can detect the majority of the metabolites. Of the ones that remain undetected, it is likely that pyruvate still elutes too early to be measured in the 4 minute cut time. One option that could be explored is the use of a column with a higher film thickness to increase the retention times. An issue with this approach is that the instrument is an institutional instrument shared with other lab groups as well as teaching courses. While this option is not impossible, it should be avoided unless it is believed that metabolites with very low molecular weight could be potential biomarkers of EC. Based on previous studies, there does not appear to be indication that this would be the case. Of the 9 metabolites previously identified as potential biomarkers, none of them are likely to elute before glycine/succinate.\(^6\)\(^{,20}\)

For the ion source temperature, the increased temperature (225 vs 200 °C) has not shown a significant difference in the data. It theoretically improves the detection of late-eluting compounds, but that has not been reflected in the data. Despite a lack of improvement, the higher temperature is preferred because it is better for long term cleanliness of the ion source and helps to reduce peak tailing.\(^{41}\) One disadvantage to using the higher source temperature is issues with spectral matching as spectra in the databases tend to be captured at lower ion source temps.\(^{41}\)

One change that was not mentioned in the method optimization was the decrease in the temperature of the transfer line from 320°C to 280°C. This change to the method was not done as a result of experimental results. During some of the analyses, there was an issue of the transfer line not being able to reach the desired temperature, and the temperature was kept in the method. The instrument has since experienced a great deal of maintenance, so the 320°C transfer line
temperature may be revisited. Perhaps the higher temperature would help with some of the late-eluting compounds.

The biggest challenge to the optimization of the instrumental method has been the fluctuation in the condition of the instrument. Ion-trap MS already has the disadvantage of having a lower sensitivity than the traditional quadrupole MS used in most biomarker studies, so it will be necessary for the ion-trap system to be at peak performance for the highest sensitivity. From the beginning, the instrument has faced large fluctuations in performance due to air leaks and contamination. When the instrument was performing well, the data collected was better as well. When the performance was low, maintenance had to be conducted to maintain decent sensitivity. Unfortunately, any maintenance that was conducted also altered the overall state of the instrument. These differences prevent all of the data that had been collected from being compared reliably. This is an important factor to consider in the clinical trial as all of the clinical samples, which is expected to total 600, must be analyzed under the exact same conditions for the biomarker results to be deemed reliable. As such, the state of the instrument will need to be monitored carefully once sample analysis begins, and any maintenance should be avoided, except for periodically changing out the septum and liner of the GC inlet.

It is suspected that many of the issues regarding the instrument’s performance are attributed to the sharing of the instrument. If other operators are not careful about the samples they inject into the GC-MS system, they could damage its performance. For example, samples that are too high in concentration could overload the detector and/or foul the inlet liner, and samples with nonvolatile or non-ionizable components could build up in the ion source. Concentration and cleanliness of the samples should be considered by all parties operating the GC-MS.
Clinical Trial Potential

Overall, the current method for the current clinical trial appears to have the potential to be a successful study once the method has been fully optimized. The methodology has generally demonstrated good reproducibility and the potential for sensitivity at low part per million levels of concentration. In testing biological samples, a large number of metabolite peaks were both unique and quantifiable. Due to the sharing of the instrument, the most important factor to the success of this trial will likely be the maintenance and monitoring of the GC-MS system prior to and throughout sample analysis. If the instrument is likely to be operated for other purposes outside of the clinical trial, other operators should be made aware of the need to be cautious regarding the concentration and cleanliness of the samples that they wish to inject into the GC-MS. It is likely that more frequent changes of the inlet liner would be required if multiple groups use the instrument. As long as this is accounted for, there should be minimal issues regarding the capability of the GC-ion trap MS system to successfully determined biomarkers for EC.

Though the potential for biomarker identification is present, it is unlikely that the obtained results will be the same as the first EC study. The results of the current clinical trial will be limited by the lowered sensitivity of the ion-trap MS, but the results should theoretically be applicable to any other MS systems with a higher sensitivity. The results of the study may also be affected in part by the choice of the multivariate classification techniques that will be used to process all of the chromatographic and spectral data.

Metabolomics Standard Initiative

An important consideration of new metabolomic investigations is the Metabolomics Standards Initiative (MSI). MSI has provided guidelines that define the various degrees of confidence with which chromatographic or spectral features are labeled. There are four MSI
levels of rigor with which metabolite identifications are recognized among the literature. MSI-Level 1 have fully identified compounds that have at least two independent measurements, such as retention time and mass spectrum matching, that match authentic standards; this is the gold standard for metabolomic research. MSI-Level 2 are putative identifications that rely on spectral matching to a database but lack comparison to an authentic standard. At most, it could be said that the current method meets the criteria for MSI-Level 2. Level 3 are putative compound class identifications, and Level 4 are completely unknown compounds, but their spectral or chromatographic features can still be statistically differentiated from the background. One of the additional goals by the end of the clinical trial is to meet MSI-Level 1 standards.

Future Work

The final method for the clinical trial is beginning to take shape, but a considerable amount of optimization is still needed. Currently, the instrumental method has been modified to provide suitable chromatographic and spectral data. As the optimization process progresses, the instrumental method may continue to be modified as needed, but the focus should now turn primarily to the sample preparation steps as changes to the instrumental method do not seem to yield significant changes to the results. Each step of sample preparation still needs to be thoroughly evaluated in terms of solvent volume, reaction time, and general reaction conditions (temperature, mixing method, etc.). So far only one of the three serum aliquots has been processed, leaving more opportunities for method development with the same serum sample.

Looking further down the line, there should be several more tests necessary following method optimization before the clinical samples are analyzed. Once the method for serum samples is optimized, the method should be tested with the other proposed biological fluids. Though the method should be applicable to saliva and urine as well, it would be best to ensure
that the same sample preparation will be suitable for the composition of these biofluids. Saliva is less of a concern as it is thought to closely resemble plasma, but urine tends to have a high concentration of urea that may need to be addressed. As further testing is conducted, samples should be randomly analyzed in larger batches that also include a blank. This will help to mimic the analysis conditions that would take place for the clinical trial. These are only a few of the suggested steps moving forward in the method development process. There are still many different options for how to continue to improve and optimize the given method.

Conclusion

While great progress has been made in developing the method and confirming the viability of the clinical trial, there are still many optimizations to be conducted before the method is ready for the clinical trial samples. A fully optimized method would ensure that the data collected will be as thorough as it can be for the given system. The success of this project would make important contributions towards identification of biomarkers for EC and the subsequent development of a screening system. Just the identification of the biomarkers alone opens up the possibility of learning about the mechanism of EC and potential treatments for the disease. The method development is a tedious and slow process, but the resulting method will be applicable to both the current study and other studies of similar nature.
References


Zullo, F.; Guida, M. Metabolomic Signature of Endometrial Cancer. J Proteome Res. 2018,
17, 804-812.
18. Bouchard-Mercier, A.; Rudkowska, I.; Lumieux, S.; Couture, P.; Vohl, M. C. The metabolic
signature associated with the Western dietary pattern: a cross-sectional study. Nutr J. 2013,
12, 158.
19. Lankadurai, B. P.; Nagato, E. G.; Simpson, M. J. Environmental metabolomics: an emerging
approach to study organism responses to environmental stressors. Environ Rev. 2013, 21,
180-205.
as revealed by ion-trap gas chromatography-mass spectrometry based metabolic profiling.
Biochim Biophys Acta. 2013, 1830, 2779-2789.
metabolomics as a novel diagnostic approach for pancreatic cancer. Metabolomics. 2010, 6,
518-528.
22. Duportet, X.; Aggio, R. B. M.; Carniero, S.; Villas-Bôas, S. G. The biological interpretation
of metabolomic data can be misled by the extraction method used. Metabolomics. 2012, 8, 3,
410-421.
23. Lu, W.; Su, X.; Klein, M. S.; Lewis, I. A.; Fiehn, O.; Rabinowitz, J. D. Metabolite
Measurement: Pitfalls to Avoid and Practices to Follow. Annu Rev Biochem. 2017, 86, 277-
304.


39. Troisi, J. University of Salerno, Salerno, Italy. Personal communication, 2019.


