

## **Gas Chromatography-Mass Spectrometric Characterization of Metabolomic Signatures for Clinical Validation of Endometrial Cancer Screening Test**

*Faculty mentor: Dr. Steven Symes*

Endometrial cancer (EC) is the most frequent malignant tumor of the female reproductive tract and the sixth most common neoplasm in women worldwide. Type I EC, which is the most frequent (80% of all cases), consists of tumors of endometrioid histology. Several risk factors for EC were reported, such as obesity, metabolic syndrome, estrogenic-only hormone after menopause, therapeutic or preventive tamoxifen, and type 2 diabetes. Also, exposure of endometrial tissue to estrogens caused by nulliparity, early menarche, late menopause, polycystic ovarian syndrome, family history of EC in a first-degree relative (mother, sister, or daughter) was reported as a risk factor. Presently, there is no effective screening system for EC, so diagnosis generally occurs only after observable cancer symptoms (like menometrorrhagia) or determination with invasive procedures (hysteroscopy or dilation and curettage).

The project will use an innovative metabolomics approach to produce a promising screening system for the early diagnosis of EC. The project will determine metabolomic signatures from 200 patients that include healthy post-menopausal women (n= 125, the 'at risk population) and post-menopausal patients (n=75) with EC. Healthy women and subjects with EC will be enrolled in a clinical trial at Erlanger Hospital through the Department of Obstetrics and Gynecology. The determination of the metabolomic signature from healthy and EC patients will be performed by means of Gas Chromatography-Mass Spectrometry (GC-MS). Serum, urine, and saliva metabolomes will be extracted and purified by means of the MetaboPrep kit.

Metabolomics is an analytical tool used in combination with bioinformatics and pattern recognition approaches to detect metabolites and follow their changes in biofluids or tissues. Metabolites are low-molecular-weight organic and inorganic chemicals that are the substrates, intermediates, and by-products of enzyme-mediated biochemical reactions in the cell. Metabolomics allows a global assessment of the cellular state, taking into account genetic regulation, altered kinetic activity of enzymes, and changes in metabolic reactions. Thus, the metabolome is considered more representative of the phenotype than the other -omic sciences, e.g., genomics, transcriptomics, and proteomics.

Metabolome extraction from multiple tissues (serum, urine, saliva) and subsequent analysis will be carried out according to standardized protocols we have used previously and from which published data resulted. The chromatographic data will be analyzed using multi-variate statistical analysis. To do this, several data pre-treatments will be used (chromatogram alignment, peak integration and normalization). Using the training data set, we will build and optimize several classification models and machine learning algorithms. All models will then be combined using a voting scheme to build an ensemble learning algorithm that uses each models' cross validation accuracy as weight for the voting scheme. This process, e.g., metabolite extraction, chromatographic analysis, building classification models and implementing machine learning, is one that we have used to build screening tests for other medical conditions (chromosomal anomalies and central nervous anomalies). We are confident that this process will also produce a successful test for endometrial cancer.