

Early-stage Detection of Ovarian Cancer through E-nose Odor Profile Determination

Jens Eriksson^{1,2} Ivan Shteplyuk¹, Lingyin Meng¹, Christer Borgfeldt³, Donatella Puglisi¹

¹ Department of Physics, Chemistry and Biology, Linköping University, 58183 Linköping, Sweden,

² VOC Diagnostics AB, LEAD Linköping AB Teknikringen 7, Sweden

³ Department of Obstetrics and Gynecology, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

jens.eriksson@liu.se

Summary:

Endogenous volatile organic compounds (VOCs) are products of metabolic activity. Metabolite concentrations are affected by tumor necrosis, resulting in cancer-specific VOC patterns. Here we introduce an e-nose for use in cancer diagnostics based on odor detection from 1 ml of blood plasma. Based on measurements of 169 samples from cancer patients (124 ovarian cancer patients and 45 endometrial cancer patients in stages I – IV) along with 161 negative controls, our machine-learning model can distinguish between cancer and healthy samples with a sensitivity of 96.9%, and a specificity of 97.1%.

Keywords: Cancer, diagnostics, electronic nose, machine learning, early detection, ovarian cancer

Background, Motivation and Objective

While the risk of dying from cancer has steadily declined over the last 30 years, global annual cancer incidences are around 19.3 million, with nearly 10 million people dying from cancer each year [1]. Many of them would not need to if the condition had been detected at an early stage. A subcategory of cancers known as the silent killers are characterized by being asymptomatic until late stages, resulting in bleaker prognosis upon diagnosis. Ovarian cancer (OC) is often asymptomatic with 60-75% of patients diagnosed at late stages [2]. To exacerbate the problem, there are also no viable screening tests available for this silent killer, meaning that early-stage detection is rare.

Description of the New Method or System

Here we introduce an electronic nose that allows shining a light on these wannabe killers while they are immature, thereby enabling their removal before they grow into killers. The primary target is early detection of OC using an approach based on condition-specific odor patterns resulting from metabolites found in blood. Endogenous volatile organic compounds (VOCs) are products of metabolic activity whose concentrations in blood are affected by tumor necrosis, resulting in VOC patterns that are specific to cancer type, subtype, and stage. The results discussed here come from measurements of 169 blood samples from cancer patients (124 OC- and 45 endometrial cancer patients in stages I – IV) together with 161 healthy controls.

The specific electronic nose [3] comprises 4 × 8 chemiresistive sensors (TGS2X series, FIGARO INC. IL, USA) designed towards interaction with different classes of molecules, arranged in four banks operating at different temperatures. Samples (blood plasma) are placed in a holder which is inserted into the instrument at the start of a measurement. After an equilibration step where VOCs are emitted into the instrument-headspace, a fan moves the VOCs over the sensors. The signals from 32 sensors upon interaction with the sample-specific VOCs are recorded for ten minutes at 10 Hz during a measurement and are used to train offline machine learning (ML) models that are used for stepwise binary classification based on 85 features extracted from pre-processed sensor data (see Fig. 1). Each classifier pertains to an ML model trained on known (labeled) data and is used to determine the class (cancer/no cancer, cancer type, and cancer stage) of blind samples. The final classification algorithm is transferred to the instrument and allows a diagnostic prediction on unknown blood samples immediately after the 10 min measurement.

Results

Signal preprocessing, feature extraction, classifier training and testing, performed using three independent approaches, all show overall accuracy higher than 98%. For the best model, binary classification as either “cancer” or “healthy” based on 85 features and a random forest

classifier resulted in sensitivity = 96.9%, specificity = 97.1%, and an ROC-AUC of 0.995 using a 10/90 train/test split. Additional classifiers were implemented sequentially for further division into cancer type and stage, allowing identifying OC and endometrial cancers with sensitivities of 94.8% and 93.1%.

Applications), funded by the Swedish Research Council, Vetenskapsrådet, grant No. 2023-07219, and Sweden's Innovation Agency, Vinnova, grant No. 2023-04186, in collaboration with the US National Science Foundation (NSF).

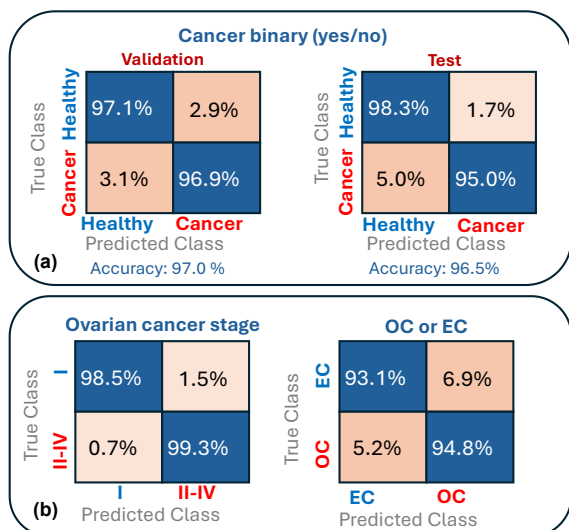


Fig. 1. Confusion matrices showing the classification at (a) validation and test stages for binary classification between healthy controls and the two studied cancer types grouped together as just one class, and (b) validation confusion matrices for classification of ovarian cancer stage I or II – IV (left panel) and distinction between ovarian- (OC) and endometrial cancer (EC) (right panel).

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomatram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021; 71: 209 – 249, <https://doi.org/10.3322/caac.21660>
- [2] Leandersson P, Hogberg T, Dickman PW, Malander S, Borgfeldt C. Incidence and survival of epithelial ovarian, fallopian tube, peritoneal, and undesignated abdominal/pelvic cancers in Sweden 1960–2014: A population-based cohort study. *BMC Cancer.* 2021;21(1):465
- [3] J. Eriksson, D. Puglisi, C. Borgfeldt, Electronic Nose for Early Diagnosis of Ovarian Cancer, *Proceedings* 97, 145 (2024); doi: 10.3390/proceedings2024097145

Acknowledgements

The work was carried out within the Strategic Innovation Programs Swelife and MedTech4Health, a joint venture of Vinnova, Formas and the Energy Agency, grant No. 2022-03464 and grant No. 2023-03874, and within the Convergence Accelerator Program (Track L – Real-World Chemical Sensing