




Article

Detection of Ovarian Cancer through Exhaled Breath by Electronic Nose: A Prospective Study

Francesco Raspagliesi ¹, Giorgio Bogani ^{1,*} , Simona Benedetti ², Silvia Grassi ² , Stefano Ferla ¹  and Susanna Buratti ²

¹ Department of Gynecologic Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, 20133 Milan, Italy; raspagliesi@istitutotumori.mi.it (F.R.); Stefano.ferla1@studenti.unimi.it (S.F.)

² Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, 20133 Milan, Italy; simona.benedetti@unimi.it (S.B.); silvia.grassi@unimi.it (S.G.); susanna.buratti@unimi.it (S.B.)

* Correspondence: giorgiobogani@yahoo.it or giorgio.bogani@istitutotumori.mi.it

Received: 12 August 2020; Accepted: 22 August 2020; Published: 25 August 2020



Abstract: Background: Diagnostic methods for the early identification of ovarian cancer (OC) represent an unmet clinical need, as no reliable diagnostic tools are available. Here, we tested the feasibility of electronic nose (e-nose), composed of ten metal oxide semiconductor (MOS) sensors, as a diagnostic tool for OC detection. Methods: Women with suspected ovarian masses and healthy subjects had volatile organic compounds analysis of the exhaled breath using e-nose. Results: E-nose analysis was performed on breath samples collected from 251 women divided into three groups: 86 OC cases, 51 benign masses, and 114 controls. Data collected were analyzed by Principal Component Analysis (PCA) and K-Nearest Neighbors' algorithm (K-NN). A first 1-K-NN (cases vs. controls) model has been developed to discriminate between OC cases and controls; the model performance tested in the prediction gave 98% of sensitivity and 95% of specificity, when the strict class prediction was applied; a second 1-K-NN (cases vs. controls + benign) model was built by grouping the non-cancer groups (controls + benign), thus considering two classes, cases and controls + benign; the model performance in the prediction was of 89% for sensitivity and 86% for specificity when the strict class prediction was applied. Conclusions: Our preliminary results suggested the potential role of e-nose for the detection of OC. Further studies aiming to test the potential adoption of e-nose in the early diagnosis of OC are needed.

Keywords: ovarian cancer; early diagnosis; electronic nose; MOS sensors; K-NN models

1. Introduction

It is estimated that over 22,000 new cases of ovarian cancer (OC) will be diagnosed and that 14,000 women will die of this disease in 2020 only in the U.S. OC is the fifth most common cancer in women and the most common cause of gynecologic cancer deaths, thus being one of the most lethal cancers in women [1]. Nearly three cases out of four were diagnosed in the advanced stage of the disease (International Federation of Gynecology and Obstetrics (FIGO) stage III or IV) at presentation [2]. The absence of valid screening programs and the rapid spread through the peritoneal surface represent the main determinants for ovarian cancer lethality. In fact, to date, no screening test is accurate and reliable enough to detect OC in women. Growing evidence suggested that clinical examination, transvaginal ultrasound, and CA-125 dosage are not enough to detect early-stage OC in the general population [3]. Recently, the Food and Drug Administration (FDA) has provided recommendations against the use of currently offered tests for OC screening, stating that these tests should not be used due to the high rate of false results [4].

Accumulating evidence has shown that analyzing the volatile organic compounds (VOC) in biological fluids such as blood, urine, and breath might provide promising biomarkers for several diseases, including solid tumors [5–7]. VOCs are gaseous molecules that can be sampled quickly and non-invasively from breath. They can originate either from within the body (endogenous VOCs) or from external sources, including medications and environmental exposure (exogenous VOCs). As other gasses (including O₂ and CO₂), VOCs pass from the bloodstream into the lungs. Nowadays, breath analyses are applied in medicine and clinical pathology as a non-invasive and comfortable way for individual health state evaluation [8–11]. Recently, e-nose has been used in breath analysis for the diagnosis of renal diseases, diabetes, asthma [8–11]. The application of the e-nose has extended in oncology, specifically in patients with lung cancer, and has also been investigated in the prostate, colonic, breast, and head and neck cancers with promising results [12–19]. Growing evidence suggests that cancer metabolism alters the volatile part of the metabolome that are present in various body fluids, including tissues, blood, and breath [20–22]. The central hypothesis of our work is that, since OC has characteristics of a systemic disease, it is possible that metabolomes produced by cancer cells, might be detected in the exhaled breath of OC patients. In the present paper, we aimed to collect preliminary data to determine whether this technology could be used as a diagnostic tool to triage patients with OC.

2. Results

Two hundred eighty-eight (288) exhaled breath samples were collected during the experimentation. Data from 37 patients were excluded since 28 samples were not analyzed on the same day of sample collection, and nine patients were excluded since they did not match inclusion criteria. Then, 251 subjects met the inclusion criteria and were included in the final analysis. The study population was divided into three main groups: (1) controls—healthy subjects without adnexal masses (114 patients; 45.4%); (2) OC cases—women with malignant adnexal masses (86 patients; 34.3%); (3) benign disease—women with suspect adnexal masses with evidence of benign disease at histological evaluation (51 patients; 20.3%). The main characteristics of the study population are displayed in Table S1. Histopathological diagnosis of malignant masses included: high-grade serous ovarian cancer ($n = 63$), low-grade serous ovarian cancer ($n = 8$), mucinous ovarian cancer ($n = 5$), ovarian carcinosarcoma ($n = 2$), endometrioid ovarian cancer ($n = 2$), and other malignant entities ($n = 6$). The whole population of OC cases included 32 (37.2%) patients with early-stage disease (i.e., stage I–II), and 54 (62.8%) patients with advanced-stage OC. Details of the histopathological findings are reported in Table S2.

2.1. E-Nose Analysis and Sensor Selection

Figure 1 shows the sensor signals collected during the sampling phase on a case (a) and a control (b) sample. The x-axis represents the sampling time, while the sensor responses are reported on the y-axis. As displayed in Figure 1, the sensor responses gradually increased during sampling, reaching a plateau after 15–20 s. From the histograms representing the responses collected after 40 s of sampling, it is possible to notice that the sensor signals were lower for the control sample, while for the OC case, three sensors (S2, S6, and S7) showed significantly increased signals. To select the MOS sensors relevant for OC cases detection, an interval plot was primarily built on the e-nose dataset transformed by column autoscaling to eliminate the numerical differences between variables. The interval plot shows the 95% confidence interval for the mean value of each group. The difference among groups is statistically relevant only if there is no overlap of the indicator bars, and the width of the intervals correlates with the amount of variation in the data.

Figure 2 shows the comparison performed on data collected by each e-nose sensor for two groups (OC cases, including early and advanced stage, and controls). A statistically significant difference between cases and controls was highlighted for eight sensors, while for two sensors (S4 and S10), the overlapping of the indicator bars showed that the two classes were not different at the given level of confidence ($\alpha: 0.05$). Furthermore, controls were characterized by small intervals showing

more consistent data and less variation, while the wide intervals of OC cases indicated a significantly higher variation.

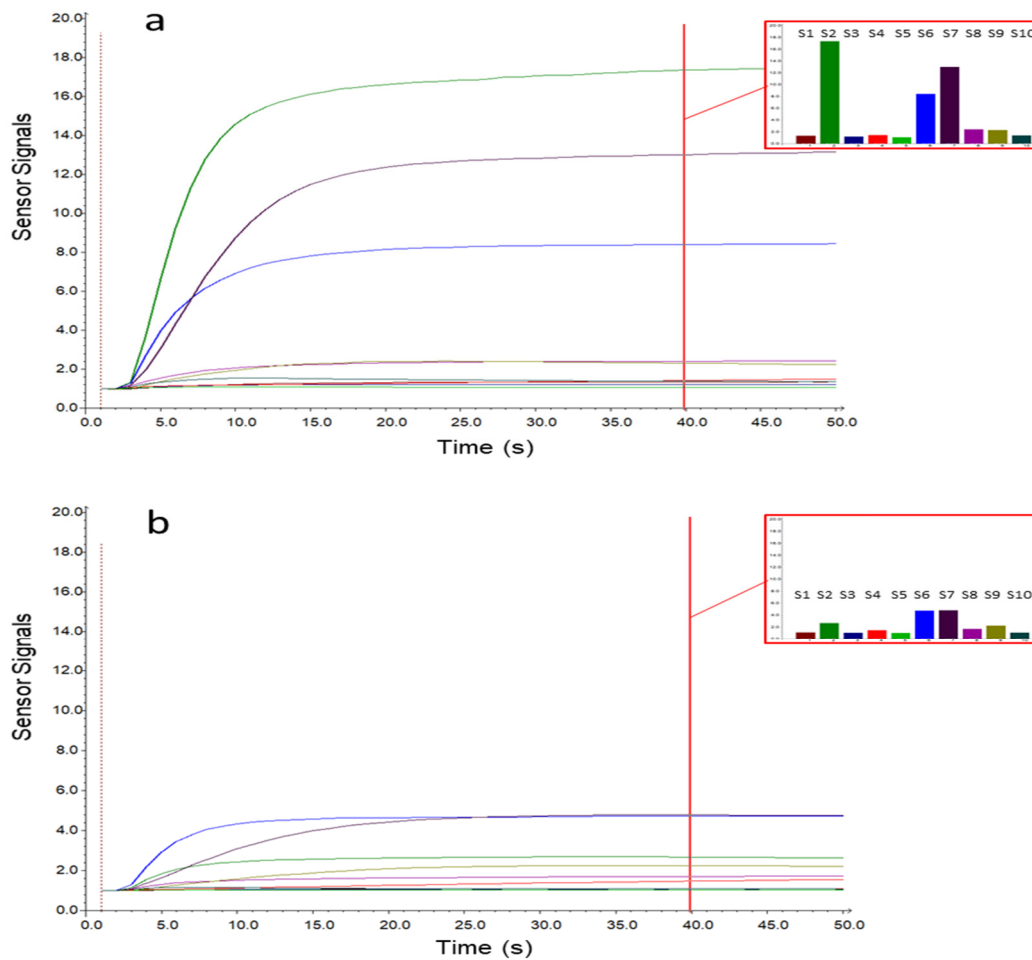


Figure 1. Metal oxide semiconductor (MOS) sensor signals for a case (a) and control (b) sample. The histograms represent the responses after 40 s of sampling for the 10 e-nose sensors (S1–10).

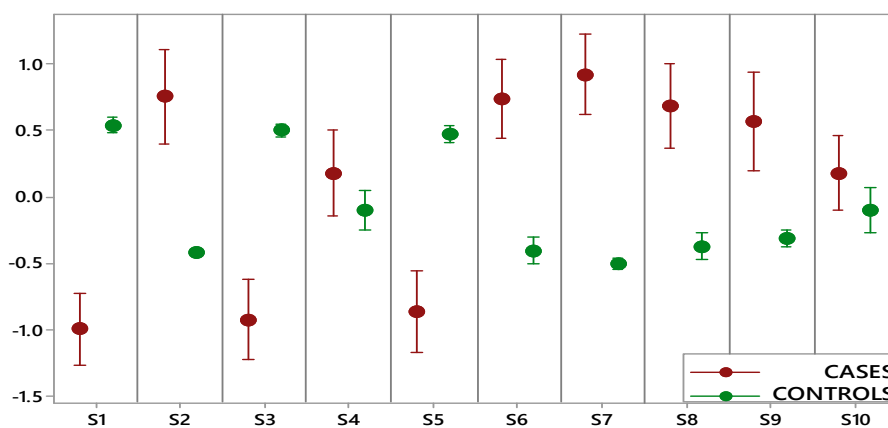


Figure 2. Interval plot of electronic nose data set for the ovarian cancer (OC) cases (red) and controls (green).

A further Principal Component Analysis (PCA) elaboration was applied to data collected by the eight selected metal oxide semiconductor (MOS) sensors, to identify the most informative sensors useful in discriminating OC cases from non-cancer samples (i.e., controls and benign disease). Figure 3 displays the PCA score plots (a1 and a2) and loading plots (b) of the data collected on OC cases (red), benign disease (yellow), and controls' (green) samples.

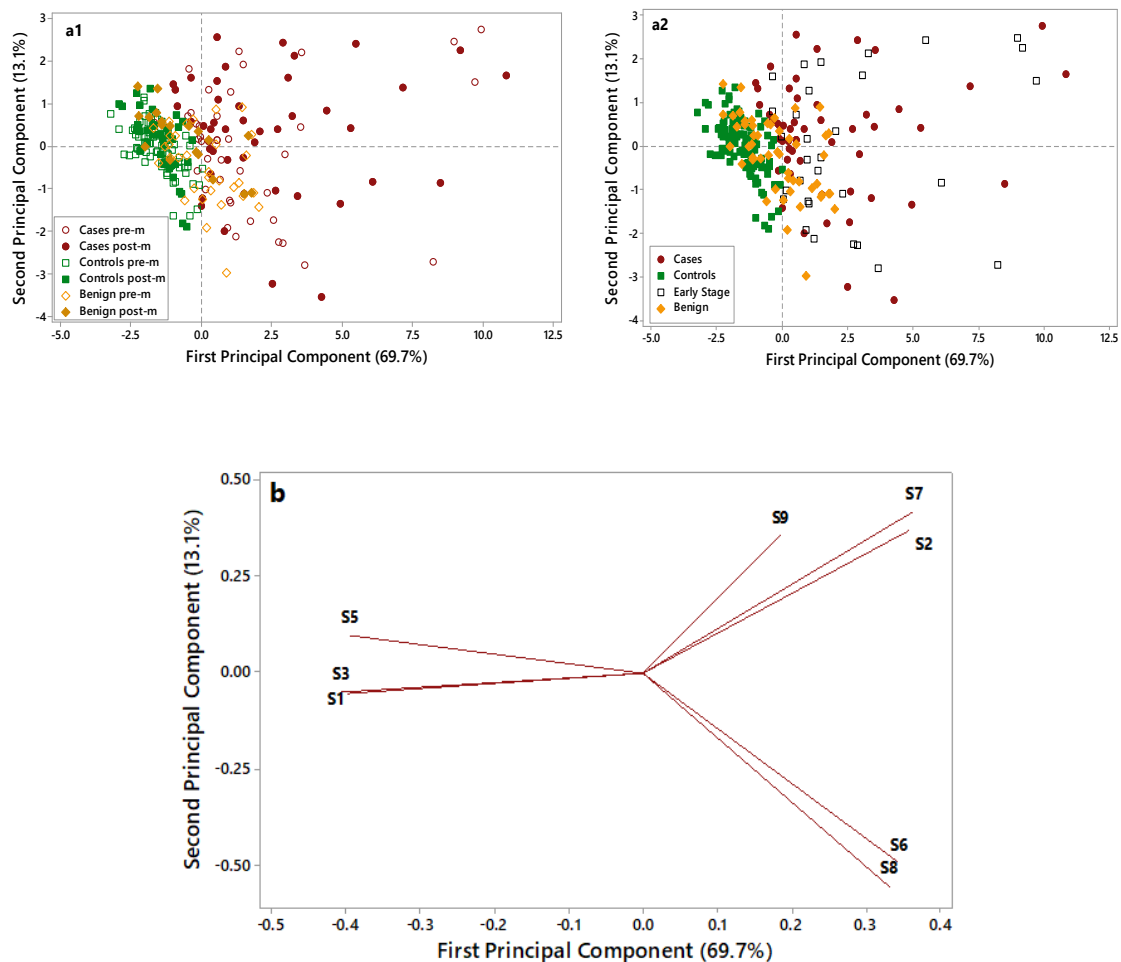


Figure 3. Principal Component Analysis(PCA) score plots (a1,a2) and loading plot (b) of electronic nose data collected on control (green), benign (yellow), and case (red) samples. In the score plot (a1) empty symbols are for premenopausal (pre-m) patients, and full symbols are for postmenopausal (post-m) patients. In the score plot (a2) empty black symbols are for early-stage patients. Two score pots were elaborated to: i. discriminate premenopausal and postmenopausal patients (Figure 3(a1)). ii. evaluate the plot distribution for early-stage ovarian cancer (OC) patients (Figure 3(a2)).

As shown in Figure 3(a1), control samples (full and empty green symbols) are grouped in the negative part of the first Principal Component (PC) and are well discriminated by OC cases (full and empty red symbols), characterized by higher PC1 scores and more dispersed along with the first and second PCs. Notably, the overlap between these two groups is very limited. Benign disease samples (full and empty yellow symbols), located in the middle of the plot, show partial overlap with both cases and controls. No grouping can be highlighted considering the pre- (empty symbols) and post (full symbols) menopausal patients.

In Figure 3(a2), early-stage patients are put in evidence (empty black symbols) to evaluate their distribution in the plot. Early-stage patients are dispersed in the right side of the plot along with cases, thus highlighting the ability of the e-nose to recognize them as cases and to discriminate early

stages from healthy controls and benign. The loading plot (Figure 3b) identified four sensors, which were more suitable to discriminate among different samples: S1 sensor was selected for its ability to characterize control samples; S2, S6, and S7 sensors were selected for their ability to discriminate OC cases through significantly increased response signals. The ability of the e-nose in identifying OC cases was irrespective from the CA125 levels (data not shown) and from the tumor size (≤ 3 / > 3 cm) (Figure S1). The e-nose was able to discriminate OC cases based on their histology (high-grade serous vs. other). Although all OC cases were clustered in the same group, cases characterized by non-serous histology were homogeneously grouped on one corner of the OC group, thus suggesting that e-nose could potentially discriminate among different histology and biology (Figure S2).

2.2. K-Nearest Neighbours Models Development

The predefined groups were used as a-priori information (Y) to build classification models based on the e-nose data collected by the four selected sensors (X). The applied classifier (KNN) was evaluated by three metrics: sensitivity, specificity, and accuracy. These were computed on the basis of four factors (Table S3): True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Two models were elaborated to discriminate between OC cases and non-cancer patients. Both models were developed using strict and most probable prediction approaches for a class assignment. A first 1-K-NN (cases vs controls) model was developed to discriminate between OC cases and controls (Table 1).

Table 1. Figures of merit for the K-Nearest Neighbors (KNN) model developed for cases vs. controls: sensitivity (Sens), specificity (Spec), and accuracy (Acc) are reported for the strict and the most probable assignation of the classes.

| Model | Calibration | | | | | Cross-Validation | | | Prediction | | | |
|--------------------------|-------------|-----|------|------|------|------------------|------|-----|------------|------|------|------|
| | Classes | N | Sens | Spec | Acc | Sens | Spec | Acc | N | Sens | Spec | Acc |
| Strict prediction | Cases | 58 | 91% | 95% | 93% | 93% | 99% | 95% | 28 | 93% | 100% | 96% |
| | Controls | 76 | 95% | 91% | 93% | 96% | 90% | 95% | 38 | 100% | 93% | 96% |
| | Global | 134 | 94% | 93% | 93% | 95% | 93% | 95% | 66 | 98% | 95% | 96% |
| Most probable prediction | Cases | 58 | 100% | 100% | 100% | 93% | 96% | 95% | 28 | 100% | 100% | 100% |
| | Controls | 76 | 100% | 100% | 100% | 96% | 93% | 95% | 38 | 100% | 100% | 100% |
| | Global | 134 | 100% | 100% | 100% | 95% | 94% | 95% | 66 | 100% | 100% | 100% |

The model was built using a calibration dataset of 58 cases and 76 controls, which was internally validated (cross-validation). The sensitivity, specificity, and accuracy of the strict class prediction approach were 95%, 93%, and 95%, respectively. Two out of the nine misclassified patients were early-stage patients, while the other 21 early-stage patients included in the cross-validation were correctly recognized as OC cases.

The remaining e-nose data (28 cases and 39 controls) were used to test the model capability in prediction, and confirmed its reliability to discriminate cases from controls (98% of sensitivity, 95% of specificity, and 96% of accuracy). Only a single early-stage patient was misclassified as a control, while the other eight early-stage patients included in the prediction test set were correctly assigned to OC cases. The most probable class assignation gave a perfect prediction of both classes, with a 100% accuracy.

A second 1-K-NN (cases vs. controls + benign) model was built by grouping the two non-cancer groups (benign disease and controls). The two classes considered in this model were: OC cases ($n = 86$) and controls ($n = 114$) plus benign disease ($n = 51$). The original dataset was split into a calibration set (70% of the data) and a test set (30% of the data). The model performance was tested in cross-validation and prediction. When the class prediction was applied, a sensitivity of 86%, a specificity of 84%, and an accuracy of 85% were achieved in cross-validation. Four out of 23 early-stage patients and two out of 14 patients with tumor size ≤ 3 cm were misclassified in the cross-validation phase. The prediction's

performance of this model had 89% sensitivity, 86% specificity, and 87% accuracy. In this validation phase, only one out of nine early-stage patients was misclassified. The most probable class assignment gave a perfect prediction of both classes, with a 100% accuracy (Table 2).

Table 2. Figures of merit for the KNN model developed for cases vs. controls + benign: sensitivity (Sens), specificity (Spec), and accuracy (Acc) are reported for the strict and the most probable assignment of the classes.

| Model | Population | | Calibration | | | Cross-Validation | | | Prediction | | | |
|--------------------------|-------------------|-----|-------------|------|------|------------------|------|-----|------------|------|------|------|
| | Classes | N | Sens | Spec | Acc | Sens | Spec | Acc | N | Sens | Spec | Acc |
| Strict prediction | cases | 58 | 79% | 89% | 84% | 81% | 89% | 85% | 28 | 82% | 93% | 87% |
| | controls | 110 | 89% | 79% | 84% | 89% | 81% | 85% | 55 | 93% | 82% | 87% |
| | global | 168 | 86% | 83% | 84% | 86% | 84% | 85% | 83 | 89% | 86% | 87% |
| Most probable prediction | cases | 58 | 100% | 100% | 100% | 81% | 89% | 85% | 28 | 100% | 100% | 100% |
| | Controls + benign | 110 | 100% | 100% | 100% | 89% | 81% | 85% | 55 | 100% | 100% | 100% |
| | global | 168 | 100% | 100% | 100% | 86% | 84% | 85% | 83 | 100% | 100% | 100% |

3. Discussion

The present study investigated the role of e-nose as a diagnostic tool for the detection of OC, thus reporting several noteworthy findings. First, e-nose could discriminate between breath samples collected from OC cases and non-cancer patients, albeit in a small and limited cohort of patients. Second, our study identifies four specific e-nose sensors, which can have a potentially relevant role in the detection of OC cases. Third, our models are characterized by high levels of sensitivity, specificity, and accuracy for the analyzed cohort of samples. Additionally, we observed that e-nose could discriminate between early-stage and advanced OC and controls, also regardless of patients' age. However, this observation was limited due to the small numbers, unequal distribution of cancer subtypes, and the low prevalence of the early-stage disease in our cohort.

The exhaled breath contains hundreds of VOCs, which reflect the metabolic processes of both normal and cancer cells, and can be detected by the e-nose technology. The feasibility of the e-nose in the diagnostic workup of several diseases has been investigated, suggesting that potentially e-nose might discriminate among types of different diseases [12–21]. There is a growing body of literature suggesting that e-nose could be considered a useful tool for the early detection of lung and head and neck cancers [15,19,20,23]. A recently published review and meta-analysis on 14 studies testing the e-nose in patients with lung cancer, suggested that this method has high (>80%) sensitivity and specificity in discriminating between patients with lung cancer and healthy subjects [9]. In the field of oncology, e-nose might be useful not only from a diagnostic point of view but also for monitoring disease response during treatment [24]. Recently, de Vries observed that the e-nose assessment was effective for the non-invasive prediction of a patient's disease response during immunotherapy [24]. In a prospective study on 143 non-small cell lung cancer patients, e-nose allowed baseline discrimination between responders and non-responders to immune checkpoint inhibitors, with potential implications in predicting response to treatment [24].

An in vitro study tested e-nose for the detection of OC [25]. This study compared samples of high-grade serous OC with samples of healthy tissues (adnexal and uterine tissues). They observed that the e-nose correctly classified 84.4% and 86.8% of cancer and healthy tissues, respectively [25]. More recently, Amal et al. performed a pilot study testing the role of e-nose in OC prediction [26]. This study analyzed samples and data of 182 women, with the following distribution among groups: 48 OC patients, 48 controls, and 86 patients with benign disease. The nanoarray output showed the potential utility of e-nose in discriminating between OC patients and controls (79% sensitivity, 100% specificity, and 89% accuracy) [26]. These preliminary results highlighted that efforts are needed

to improve the sensitivity and accuracy of the e-nose. Here, we developed two models that could improve the diagnostic value of e-nose.

Screening and diagnostic methods for early diagnosis of OC represent an unmet clinical need, as no reliable screening and diagnostic devices are available. Since the diagnosis of OC is challenging, investigation of new a highly sensitive (at least 80%) and specific (as close as possible to 100%) test is paramount. We auspicate that the adoption of e-nose (in addition to conventional tests) would be useful in triaging patients needing further diagnostic and therapeutic procedures. Based on the results of the present exploratory study, the proposed approach seems to be promising; but it needs further validation. Besides, even in the presence of benign gynecologic disease, the 1K-NN model approximates quite satisfactorily the target values established for specificity and sensitivity when a strict prediction approach was applied, while for the most probable prediction approach, the target values are satisfied.

In the present paper, we performed a preliminary analysis evaluating the use of the e-nose in detecting OC. Our paper highlighted that the specificity in discriminating between OC and benign lesions was 86%. Although specificity is not very high, it is similar to those achieved with other available models. In particular, the IOTA group (International Ovarian Tumor Analysis) elaborated models to identify early-stage OC. In a recent paper, the IOTA group compared the performance of the Simple Rules Risk (SRR), and the Assessment of Different NEoplasias in the adneXa (ADNEX) model. This model was based on ultrasound, CA125 levels, and several other anamnestic features. At a 1% risk cut-off, specificity for SRR was 38.0% (35.5% to 40.6%), and for ADNEX was 19.4% (17.4% to 21.5%). At a 30% risk cut-off, specificity for SRR was 81.1% (79% to 83%), and for ADNEX was 84.5% (82.6% to 86.3%). The specificity reported with e-nose is higher than the specificity provided with these models, which is quite low [27]. Moreover, as evidenced in Table 2, the most probable class assignation gave a perfect prediction of both classes leading to 100% sensitivity and specificity.

E-nose achieved higher specificity values than those obtained based on a combination of CA125 levels and ultrasound. It is not surprising, since ultrasound and CA125 are not characterized with high specificity. Moreover, we can speculate that a new diagnostic algorithm, incorporating e-nose data with anamnestic, clinical, radiological (i.e., ultrasound) and humoral (i.e., CA125) features, would improve the specificity of this test that allows us to identify patients with OC in the early phase of the disease. The main strength of the present investigation includes the prospective evaluation of consecutive women diagnosed with suspected ovarian masses (including OC and benign ovarian masses) and healthy subjects. Other strengths of the study are: (i) the possible application of a simplified e-nose (based on the four selected sensors) that represents an innovative and promising tool for screening, diagnosis, and prognosis of patients with OC; and (ii) the development of two mathematical models that might be applied to simplify the interpretation of the results provided by e-nose.

Six points of our study deserve to be addressed. First, we observed that altered VOCs is a symptom of OC; therefore, the next step will be to investigate in a large population if altered VOCs might be predictive for higher risk of developing OC in comparison to normal VOCs, allowing to identify patients at risk, regardless the presence of ovarian masses. Second, the presence of different types of OC may have influenced the findings of the current study; future studies should stratify the results according to disease histology to assess potential differences in the e-nose detection accuracy. Third, the impact of other tumors (e.g., lung cancer) on the VOCs has to be tested to investigate the potential confounding effects on VOCs. Fourth, the inherent biases related to the relatively small sample size requires further studies to confirm the clear efficacy profile of e-nose. Fifth, further evidence is needed to assess the role of e-nose for screening and diagnostic purposes. Finally, our study suggested that the size of the tumor (<3cm vs. >3 cm) did not impact on the diagnostic ability of e-nose. However, further investigations are needed to assess the possible existence of the “lower detectable volume” and to clarify the effective diagnostic value of e-nose in the early phase of the disease. It is important to highlight that our results should be considered preliminary. This is a preliminary experience to determine whether this technology could successfully triage patients with OC. Further evidence is

warranted to evaluate the real advantages of adopting e-nose and how to improve the specificity of this approach.

4. Methods

4.1. Ethical Approval and Recruitment of Subjects

This is a case-control prospective study. The Institutional Review Board (IRB) approved this study (INT-D237354: 011/11). The study population included all consecutive patients with adnexal masses treated at the National Cancer Institute (Milan - Italy) from 1 March 2018 to 30 April 2019. All the patients signed written consent for research purposes. The exclusion criteria were: age < 18 years, presence of synchronous solid cancer (within five years), surgery for recurrent disease, and consent withdrawn. All consecutive patients diagnosed with adnexal masses scheduled to have surgery were included in the study. Additionally, we included a control group of healthy subjects without ovarian abnormalities. All the controls had a negative ultrasonographic examination. According to the study protocol, all exhaled breaths were analyzed on the same day of sample collection. OC patients included in the study had staging surgery or cytoreduction according to the disease characteristics. Details regarding our surgical practice were reported elsewhere [28,29]. The stage of the disease and grading were assessed using the International Federation of Obstetrics and Gynecologists (FIGO) system. Histological subtypes were reported according to the World Health Organization (WHO) system [29]. Patients with adnexal masses were tested the day before surgery, while healthy controls were tested at their convenience. No patient was under active oncologic treatments. The present research follows both the STARD and Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines for a study investigating diagnostic accuracy [30,31]. The objective of the adoption of the STARD and TRIPOD guidelines is to improve the completeness and transparency of the reporting of studies of diagnostic accuracy, to allow readers to assess the potential for bias in the study (internal validity), and to evaluate its generalizability (external validity) [30,31].

4.2. Sample Collection

The sample collection started without specific conditions since 2016. A preliminary feasibility study was conducted collecting data from 150 patients. The evaluation of these preliminary data allowed the identification of the correct sampling condition and its reliability (data regarding our preliminary experience are reported in Table S1). For the present work, e-nose data were collected from 1 March 2018, to 30 April 2019. During this period, exhaled breath samples were collected between 7.00 and 7.30 a.m. from fasting patients (i.e., patients were not allowed to take any food and beverages, except water, since midnight before sampling. Similarly, smoking was not allowed since the night before sampling. An accurate description of the method used for the collection of exhaled breath is reported elsewhere [17]. Briefly, subjects breathed through a Teflon mouthpiece in a homemade Nalophan bag (30 × 20 cm) taking care to evaluate alveolar breath that refers to the last portion of the exhaled breath, expelled from within the lungs and the lower-airways, which have undergone gaseous exchange with the blood in the alveoli. A healthy adult expires approximately 500 mL air with each breath, of which the first 150 mL consists of dead-space air (no transfer of oxygen) from the upper air ways and nasopharynx. Therefore, patients and controls were asked to perform a single slow vital capacity breath, to trap the last 350 mL of exhaled breath. Breath samples were analyzed by e-nose on the same collection day.

4.3. Electronic Nose Analysis

Measurements of breath samples were performed by a commercial and portable e-nose (PEN 3, Win Muster Airsens Analytic Inc., Schwerin, Germany). The sensor array is composed of 10 Metal Oxide Semiconductor (MOS) sensors of different chemical composition, thickness and working temperature, to provide sensitivity and selectivity towards volatile compounds as indicated by the instrument

supplier: S1 (aromatic compounds), S2 (broad-range compounds, polar compounds, nitrogen oxides, and ozone), S3 (ammonia, aromatic compounds, aldehydes, and ketones), S4 (hydrogen), S5 (alkanes, aromatic compounds, and less polar compounds), S6 (methane and broad-range compounds), S7 (sulfur compounds, terpenes, and sulfur organic compounds), S8 (alcohols, partially aromatic compounds, and ketones), S9 (aromatic compounds and sulfur organic compounds), and S10 (methane). During the sampling procedure, the breath samples contained in the bags were pumped over the sensor surfaces for 50 seconds (sampling time) during which the sensor responses were recorded; the flow rate was set at 400 mL/minutes. After sample analysis, the sensors were purged for 180 seconds with filtered air (purging time), then, before the next sample injection, the sensor baselines were re-established for 5 seconds. The sensor response corresponded to the fractional value obtained by subtracting the resistance signal of the baseline (R_0 -ohm) from the resistance signal of the sensors (R -ohm) and dividing by the resistance signal of the baseline (R_0), thus providing a dimensionless, normalized response. Data were acquired after 40 seconds of sampling. Each breath sample was analyzed in duplicate, and the average of the sensor data was used for the statistical elaboration. Biostatisticians who analyzed e-nose data were blinded (i.e., not aware of the patients' diagnoses) when performing the analyses.

4.4. Statistical Analysis

The e-nose data were transformed by column autoscaling and then explored by basic descriptive statistics and by multivariate statistical techniques as Principal Component Analysis (PCA) and the K-Nearest Neighbors' algorithm (K-NN). PCA is an unsupervised exploratory procedure that allows us to explore in a reduced space the data structure and the relationships between objects and variables, thanks to graphical outputs (i.e., score plot, loading plot, and bi-plot) [32,33]. To develop the classification models, the K-Nearest Neighbors' algorithm (K-NN) was used. K-NN is a simple non-linear classification approach based on Euclidean distance; it is one of the less prone to overfitting thanks to the simplicity of its algorithm, which does not include the optimization of numerous parameters and does not require any assumptions on the underlying data distribution. In detail, K-NN predicts the class membership of a sample based on the class of the K nearest sample(s) in the multidimensional space; in this work, a K value of 1 was applied. Before the development of the classification models, each dataset was split by the Kennard–Stone algorithm [32–34] into a training and a test set composed of about 70% and 30% of the original samples, respectively. The training dataset was used for the classification rule development, and the internal validation by cross-validation with five cancellation groups, whereas the test set was used to test the model's performance in the prediction. The class predictions are based on predicted class probabilities for samples; different approaches can be applied to assign a sample to a specific class. In this work we applied two strategies, namely strict class prediction and most probable prediction. Strict class prediction assigns each sample to a specific class if the belonging probability is greater than a specified threshold probability value ($p > 0.5$) for one and only one class. This approach leads to the possibility of unassigned or multiple assignments of the samples. Indeed, if a sample has a probability lower than the threshold for the two classes, no class could be assigned to the sample. On the contrary, if a sample exceeds the threshold in the two considered classes, it will be assigned to both. This prediction provides the most safety in-class assignment but could lead to poorer figures of merits. Moreover, the most probable approach has been applied to assign samples to classes. In this case, samples are assigned to the class reaching the highest probability regardless of the magnitude of that probability. Even if more than one class has a probability higher than 0.50, the sample will be assigned to the one with the highest probability; likewise, if all probabilities are below 0.50, the sample will be assigned to the one with the largest probability. This approach is advantageous if there is a need for a class assignment and if "no class" or "multiple class" has no meaning. In our case, this approach is adequate since we must assign samples in the predefined classes.

For both approaches, the classification performance was evaluated by considering the accuracy, corresponding to the probability of doing a correct classification; the sensitivity, corresponding to the probability to classify a subject as a case when this is true (true positive rate); the specificity, corresponding to the probability of classifying a subject as control when this is true (true negative rate). All the data analyses were performed under the Matlab environment (R2017b, The Mathworks Inc., Natick, MA, USA), eventually using the PLS toolbox (ver. 8.5, Eigenvector Research Inc., Manson, WA, USA) software package.

5. Conclusions

Our data suggest that e-nose is a non-invasive, appealing technique that might be useful during an OC workup. To date, our data should be considered preliminary. They represent the basis for further studies. Further large prospective studies are needed to confirm our data. OC represents a group of heterogeneous diseases [35–40]; therefore, further investigations are needed to evaluate the applicability of e-nose in detecting various OC types. The new perspective is the implementation of this device in the new collection of exhaled breath samples on a larger population to validate these results and build more robust classification models. Additionally, further investigations would evaluate the role of e-nose alone and in combination with other diagnostic tools, such as the ROMA and ADNEX models.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/12/9/2408/s1>, Figure S1: The impact of tumor size on e-nose signals, Figure S2: The impact of histology on e-nose signals, Table S1: Baseline patients' characteristics, Table S2: Final histological diagnosis for patients with adnexal masses who had surgery, Table S3: Generic confusion matrix identifying true positive (TP), false positive (FP), true negative (TN), false negative (FN) and the reference equation to calculate sensitivity (SENS) and specificity (SPEC).

Author Contributions: Conceptualization: F.R., Methodology: S.B. (Simona Benedetti), S.G., S.B. (Susanna Buratti); Data Collection: G.B., S.F.; Project administration: F.R.; Supervision: S.B. (Simona Benedetti), F.R.; writing—original draft: G.B., S.B. (Simona Benedetti); writing—review & editing: all authors. All authors have read and agreed to the published version of the manuscript.

Funding: No funding sources supported this investigation.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

| | |
|------|---|
| OC | ovarian cancer |
| MOS | metal oxide semiconductor |
| PCA | Principal component analysis |
| K-NN | K-Nearest Neighbors' algorithm |
| VOC | volatile organic compounds |
| FIGO | International Federation of Gynecology and Obstetrics |
| FDA | Food and Drug Administration |
| IRB | Institutional Review Board |
| WHO | World Health Organization |
| PC | Principal component |
| TP | true positive |
| FP | false positive |
| TN | true negative |
| FN | false negative |

References

1. Siegel, R.L.; Mph, K.D.M.; Jemal, A. Cancer statistics, 2020. *CA A Cancer J. Clin.* **2020**, *70*, 7–30. [[CrossRef](#)] [[PubMed](#)]

2. Bogani, G.; Ditto, A.; Pinelli, C.; Lopez, S.; Chiappa, V.; Raspagliesi, F. Ten-year follow-up study of long-term outcomes after conservative surgery for early-stage ovarian cancer. *Int. J. Gynaecol. Obstet.* **2020**, *150*, 169–176. [[CrossRef](#)] [[PubMed](#)]
3. Pinsky, P.F.; Yu, K.; Kramer, B.S.; Black, A.; Buys, S.S.; Partridge, E.; Gohagan, J.; Berg, C.D.; Prorok, P.C. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15years follow-up. *Gynecol. Oncol.* **2016**, *143*, 270–275. [[CrossRef](#)] [[PubMed](#)]
4. FDA. The FDA Recommends against Using Screening Tests for Ovarian Cancer Screening: FDA Safety Communication. 2016. Available online: <http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/ucm519413.htm> (accessed on 11 August 2020).
5. Di Lena, M.; Porcelli, F.; Altomare, D.F. Volatile organic compounds as new biomarkers for colorectal cancer: A review. *Color. Dis.* **2016**, *18*, 654–663. [[CrossRef](#)] [[PubMed](#)]
6. Amal, H.; Leja, M.; Funka, K.; Lasina, I.; Skapars, R.; Siviņš, A.; Ancāns, G.; Kikuste, I.; Vanags, A.; Tolmanis, I.; et al. Breath testing as potential colorectal cancer screening tool. *Int. J. Cancer* **2016**, *138*, 229–236. [[CrossRef](#)] [[PubMed](#)]
7. Khan, O.F.; Zaia, E.W.; Jhunjhunwala, S.; Xue, W.; Cai, W.; Yun, D.S.; Barnés, C.M.; Dahlman, J.E.; Dong, Y.; Pelet, J.M.; et al. Dendrimer-Inspired Nanomaterials for the in Vivo Delivery of siRNA to Lung Vasculature. *Nano Lett.* **2015**, *15*, 3008–3016. [[CrossRef](#)] [[PubMed](#)]
8. Hashim, D.; Genden, E.; Posner, M.; Hashibe, M.; Boffetta, P. Head and neck cancer prevention: From primary prevention to impact of clinicians on reducing burden. *Ann. Oncol.* **2019**, *30*, 744–756. [[CrossRef](#)]
9. Swanson, B.; Fogg, L.; Julion, W.; Arrieta, M.T. Electronic Nose Analysis of Exhaled Breath Volatiles to Identify Lung Cancer Cases. *J. Assoc. Nurses AIDS Care* **2020**, *31*, 71–79. [[CrossRef](#)]
10. Van Keulen, K.E.; Jansen, M.E.; Schrauwen, R.W.M.; Kolkman, J.J.; Siersema, P.D. Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer. *Aliment. Pharmacol. Ther.* **2019**, *51*, 334–346. [[CrossRef](#)]
11. Wilson, A.D.; Baietto, M. Applications and Advances in Electronic-Nose Technologies. *Sensors* **2009**, *9*, 5099–5148. [[CrossRef](#)]
12. Lin, Y.-J.; Guo, H.-R.; Chang, Y.-H.; Kao, M.-T.; Wang, H.-H.; Hong, R.-I. Application of the electronic nose for uremia diagnosis. *Sens. Actuators B Chem.* **2001**, *76*, 177–180. [[CrossRef](#)]
13. Yu, J.B.; Byun, H.G.; So, M.S.; Huh, J.S. Analysis of diabetic patient's breath with conducting polymer sensor array. *Sens. Actuators B Chem.* **2005**, *108*, 305–308. [[CrossRef](#)]
14. Dragonieri, S.; Schot, R.; Mertens, B.J.; Le Cessie, S.; Gauw, S.A.; Spanevello, A.; Resta, O.; Willard, N.P.; Vink, T.J.; Rabe, K.F.; et al. An electronic nose in the discrimination of patients with asthma and controls. *J. Allergy Clin. Immunol.* **2007**, *120*, 856–862. [[CrossRef](#)] [[PubMed](#)]
15. Di Natale, C.; Macagnano, A.; Martinelli, E.; Paolesse, R.; D'Arcangelo, G.; Roscioni, C.; Finazzi-Agrò, A.; D'Amico, A. Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors. *Biosens. Bioelectron.* **2003**, *18*, 1209–1218. [[CrossRef](#)]
16. D'Amico, A.; Santonico, M.; Pennazza, G.; Capuano, R.; Vespasiani, G.; Del Fabbro, D.; Paolesse, R.; Di Natale, C.; Martinelli, E.; Agrò, E.F. A Novel Approach for Prostate Cancer Diagnosis using a Gas Sensor Array. *Procedia Eng.* **2012**, *47*, 1113–1116. [[CrossRef](#)]
17. De Meij, T.G.; Ben Larbi, I.; Van Der Schee, M.P.; Lentferink, Y.E.; Paff, T.; Droste, J.S.T.S.; Mulder, C.J.; Van Bodegraven, A.A.; De Boer, N.K. Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: Proof of principle study. *Int. J. Cancer* **2014**, *134*, 1132–1138. [[CrossRef](#)]
18. Lavra, L.; Catini, A.; Olivieri, A.; Capuano, R.; Salehi, L.B.; Sciacchitano, S.; Bartolazzi, A.; Nardis, S.; Paolesse, R.; Martinelli, E.; et al. Investigation of VOCs associated with different characteristics of breast cancer cells. *Sci. Rep.* **2015**, *5*, 13246. [[CrossRef](#)]
19. Leunis, N.; Boumans, M.-L.; Kremer, B.; Din, S.; Stobberingh, E.; Kessels, A.G.H.; Kross, K.W. Application of an electronic nose in the diagnosis of head and neck cancer. *Laryngoscope* **2014**, *124*, 1377–1381. [[CrossRef](#)]
20. Van De Goor, R.M.G.E.; Leunis, N.; Van Hooren, M.R.A.; Francisca, E.; Masclee, A.; Kremer, B.; Kross, K.W. Feasibility of electronic nose technology for discriminating between head and neck, bladder, and colon carcinomas. *Eur. Arch. Oto-Rhino-Laryngol.* **2016**, *274*, 1053–1060. [[CrossRef](#)]

21. Chilo, J.; Horvath, G.; Lindblad, T.; Olsson, R. Electronic Nose Ovarian Carcinoma Diagnosis Based on Machine Learning Algorithms. In *Industrial Conference on Data Mining*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 13–23.
22. Horvath, G.; Andersson, H.; Paulsson, G. Characteristic odour in the blood reveals ovarian carcinoma. *BMC Cancer* **2010**, *10*, 643. [[CrossRef](#)]
23. Abbasi, J. “Electronic Nose” Predicts Immunotherapy Response. *JAMA* **2019**, *322*, 1756. [[CrossRef](#)] [[PubMed](#)]
24. De Vries, R.; Muller, M.; Van Der Noort, V.; Theelen, W.; Schouten, R.; Hummelink, K.; Muller, S.; Wolf-Lansdorf, M.; Dagelet, J.; Monkhorst, K.; et al. Prediction of response to anti-PD-1 therapy in patients with non-small-cell lung cancer by electronic nose analysis of exhaled breath. *Ann. Oncol.* **2019**, *30*, 1660–1666. [[CrossRef](#)] [[PubMed](#)]
25. Atalay, C.; Dogan, L.; Atalay, G. Anti-CENP-B antibodies are associated with prolonged survival in breast cancer. *Future Oncol.* **2010**, *6*, 471–477. [[CrossRef](#)] [[PubMed](#)]
26. Amal, H.; Shi, D.-Y.; Ionescu, R.; Zhang, W.; Hua, Q.-L.; Pan, Y.-Y.; Tao, L.; Liu, H.; Haick, H. Assessment of ovarian cancer conditions from exhaled breath. *Int. J. Cancer* **2015**, *136*, E614–E622. [[CrossRef](#)] [[PubMed](#)]
27. Froyman, W.; Wynants, L.; Landolfo, C.; Bourne, T.; Valentin, L.; Testa, A.C.; Sladkevicius, P.; Franchi, D.; Fischerová, D.; Savelli, L.; et al. Validation of the Performance of International Ovarian Tumor Analysis (IOTA) Methods in the Diagnosis of Early Stage Ovarian Cancer in a Non-Screening Population. *Diagnostics* **2017**, *7*, 32. [[CrossRef](#)] [[PubMed](#)]
28. Bogani, G.; Matteucci, L.; Tamberi, S.; Arcangeli, V.; Ditto, A.; Maltese, G.; Signorelli, M.; Martinelli, F.; Chiappa, V.; Maggiore, U.L.R.; et al. The Impact of Number of Cycles of Neoadjuvant Chemotherapy on Survival of Patients Undergoing Interval Debulking Surgery for Stage IIIC-IV Unresectable Ovarian Cancer: Results from a Multi-Institutional Study. *Int. J. Gynecol. Cancer* **2017**, *27*, 1856–1862. [[CrossRef](#)]
29. Raspagliesi, F.; Bogani, G.; Ditto, A.; Martinelli, F.; Chiappa, V.; Borghi, C.; Scaffa, C.; Morano, F.; Maltese, G.; Lorusso, D. Implementation of Extensive Cytoreduction Resulted in Improved Survival Outcomes for Patients with Newly Diagnosed Advanced-Stage Ovarian, Tubal, and Peritoneal Cancers. *Ann. Surg. Oncol.* **2017**, *24*, 3396–3405. [[CrossRef](#)]
30. Cohen, J.F.; Korevaar, D.A.; Altman, D.G.; Bruns, D.E.; Gatsonis, C.A.; Hooft, L.; Irwig, L.; Levine, D.; Reitsma, J.B.; De Vet, H.C.W.; et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* **2016**, *6*, e012799. [[CrossRef](#)]
31. Moons, K.; Altman, U.G.; Reitsma, J.B.; Ioannidis, J.P.; Macaskill, P.; Steyerberg, E.W.; Vickers, A.J.; Ransohoff, D.F.; Collins, G.S. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann. Intern. Med.* **2015**, *162*, W1–W73. [[CrossRef](#)]
32. Jolliffe, I. Principal Component Analysis. In *Wiley StatsRef: Statistics Reference Online*; Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L., Eds.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2014. [[CrossRef](#)]
33. Malegori, C.; Grassi, S.; Ohm, J.-B.; Anderson, J.; Marti, A. GlutoPeak profile analysis for wheat classification: Skipping the refinement process. *J. Cereal Sci.* **2018**, *79*, 73–79. [[CrossRef](#)]
34. Grassi, S.; Vitale, R.; Alamprese, C. An exploratory study for the technological classification of egg white powders based on infrared spectroscopy. *LWT* **2018**, *96*, 469–475. [[CrossRef](#)]
35. Dugo, M.; Devecchi, A.; De Cecco, L.; Cecchin, E.; Mezzanzanica, D.; Sensi, M.L.; Bagnoli, M. Focal Recurrent Copy Number Alterations Characterize Disease Relapse in High Grade Serous Ovarian Cancer Patients with Good Clinical Prognosis: A Pilot Study. *Genes* **2019**, *10*, 678. [[CrossRef](#)] [[PubMed](#)]
36. Arason, A.; Agnarsson, B.A.; Johannsdottir, G.; Johannsson, O.T.; Hilmarsdottir, B.; Reynisdottir, I.; Barkardottir, R.B. The BRCA1 c.4096+3A>G Variant Displays Classical Characteristics of Pathogenic BRCA1 Mutations in Hereditary Breast and Ovarian Cancers, But Still Allows Homozygous Viability. *Genes* **2019**, *10*, 882. [[CrossRef](#)] [[PubMed](#)]
37. Yang, W.; Shin, H.-Y.; Cho, H.; Chung, J.-Y.; Lee, E.-J.; Kim, J.-H.; Kang, E.-S. TOM40 Inhibits Ovarian Cancer Cell Growth by Modulating Mitochondrial Function Including Intracellular ATP and ROS Levels. *Cancers* **2020**, *12*, 1329. [[CrossRef](#)]

38. Sari, A.N.; Bhargava, P.; Dhanjal, J.K.; Putri, J.F.; Radhakrishnan, N.; Shefrin, S.; Ishida, Y.; Terao, K.; Sundar, D.; Kaul, S.C.; et al. Combination of Withaferin-A and CAPE Provides Superior Anticancer Potency: Bioinformatics and Experimental Evidence to Their Molecular Targets and Mechanism of Action. *Cancers* **2020**, *12*, 1160. [[CrossRef](#)]
39. Goyeneche, A.; Lisio, M.-A.; Fu, L.; Srinivasan, R.; Capuccino, J.V.; Gao, Z.-H.; Telleria, C.M. The Capacity of High-Grade Serous Ovarian Cancer Cells to Form Multicellular Structures Spontaneously along Disease Progression Correlates with Their Orthotopic Tumorigenicity in Immunosuppressed Mice. *Cancers* **2020**, *12*, 699. [[CrossRef](#)]
40. Affatato, R.; Carrassa, L.; Chilà, R.; Lupi, M.; Restelli, V.; Damia, G. Identification of PLK1 as a New Therapeutic Target in Mucinous Ovarian Carcinoma. *Cancers* **2020**, *12*, 672. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).