

Article

A Rapid and Affordable Screening Tool for Early-Stage Ovarian Cancer Detection Based on MALDI-ToF MS of Blood Serum

Ricardo J. Pais ^{1,2} , Raminta Zmuidinaite ³ , Jonathan C. Lacey ³, Christian S. Jardine ³ and Ray K. Iles ^{3,4,*}
¹ Bioenhancer Systems, Office 63 182-184 High Street North, East Ham, London E6 2JA, UK; rjpais@bioenhancersystems.com

² Centro de investigação Interdisciplinar Egas Moniz (CiiEM), Instituto Universitário Egas Moniz, 2829-511 Caparica, Portugal

³ MAP Sciences Ltd., The iLab, Stannard Way, Priory Business Park, Bedford MK44 3RZ, UK; raminta.zmuidinaite@mapsciences.com (R.Z.); jon.lacey@mapsciences.com (J.C.L.); christian.jardine@mapsciences.com (C.S.J.)

⁴ NISAD (Lund), Medicon Village, SE-223 81 Lund, Sweden

* Correspondence: ray.iles@mapsciences.com

Abstract: Ovarian cancer is a worldwide health issue that grows at a rate of almost 250,000 new cases every year. Its early detection is key for a good prognosis and even curative surgery. However, current medical examination methods and tests have been inefficient in detecting ovarian cancer at the early stage, leading to preventable death. So far, new screening tests based on molecular biomarker analysis techniques have not resulted in any substantial improvement in early-stage diagnosis and increased survival. Thus, whilst there remains clear potential to improve outcomes through early detection, novel approaches are needed. Here, we postulated that MALDI-ToF-mass-spectrometry-based tests can be a solution for effective screening of ovarian cancer. In this retrospective cohort study, we generated and analyzed the mass spectra of 181 serum samples of women with and without ovarian cancer. Using bioinformatics pipelines for analysis, including predictive modeling and machine learning, we found distinct mass spectral patterns composed of 9–20 key combinations of peak intensity or peak enrichment features for each stage of ovarian cancer. Based on a scoring algorithm and obtained patterns, the optimal sensitivity for detecting each stage of cancer was 95–97% with a specificity of 97%. Scoring all algorithms simultaneously could detect all stages of ovarian cancer at 99% sensitivity and 92% specificity. The results further demonstrate that the matrix and mass range analyzed played a key role in improving the mass spectral data quality and diagnostic power. Altogether, with the results reported here and increasing evidence of the MS assay's diagnostic accuracy and instrument robustness, it has become imminent to consider MS in the clinical application for ovarian cancer screening.

Keywords: mass spectrometry; MALDI-ToF; ovarian cancer; bioinformatics; machine learning; secretome; bioinformatics



Citation: Pais, R.J.; Zmuidinaite, R.; Lacey, J.C.; Jardine, C.S.; Iles, R.K. A Rapid and Affordable Screening Tool for Early-Stage Ovarian Cancer Detection Based on MALDI-ToF MS of Blood Serum. *Appl. Sci.* **2022**, *12*, 3030. <https://doi.org/10.3390/app12063030>

Academic Editors: Enrico Dávoli and Sen Takeda

Received: 9 February 2022

Accepted: 11 March 2022

Published: 16 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ovarian cancer is the fifth most common cause of death in women and the leading cause of gynecological-related mortality in the world [1–3]. The incidence of ovarian cancer is growing at a rate of almost 250 thousand new cases every year and may further increase in the future as populations age [2–5]. Unfortunately, the diagnosis of most cases occurs when the cancer is already too advanced for surgery alone to be curative (stages 3 and 4) [2–4]. At these stages, the cancer has already invaded widely and metastasized, thus accounting for the poor prognosis statistics associated with this disease [1–3]. For these reasons, the detection of ovarian cancer should be at its early stage (stages 1 or 2) where surgery alone or in combination with other modalities can be curative [3,6]. Therefore, the focus of health care systems has been directed towards screening programs of at-risk

female populations [3,7]. As this type of cancer shows no symptoms during initial stages, the early detection must rely on the least invasive and the most reliable screening tests to prevent potential damage to women's reproductive systems due to further invasive examination [2,4,5].

Currently, ovarian cancer screening programs are based on annual medical check-ups consisting of pelvic examination combined with a transvaginal ultrasound (TVUS) and serum CA-125 measurement by immunoassay [3,5]. TVUS imaging of the uterus, fallopian tubes, and ovaries for unusual masses detects benign and malignant masses with poor discrimination [2,5]. The CA-125 blood test has been used clinically for more than 30 years and is based on quantification of the CA-125 mucin-glycoprotein epitope in the blood, using a threshold cut-off value as a biomarker for ovarian cancer and other malignancies [3,5]. However, this biomarker fails to detect 50% of ovarian cancer at its early stage (FIGO/AJCC stage 1 and 2), 25% at stage 3, and 10% at stage 4, motivating researchers to search for new biomarkers [3,5]. Furthermore, the current costs of these screening programs are a huge burden for health care systems, leading many in the USA to stop supporting these programs due to the low accuracy for early stage detection [2,7]. Recently, some tests based on combining multiple novel blood biomarkers and circulating tumor-derived RNA signatures have shown higher sensitivities and have been proposed as potential future solutions [2,8,9]. However, these approaches are also associated with a substantial increase in the costs per patient, which will make it difficult for many countries to implement screening programs for their general population. Based on these practical considerations, it is urgent to find novel and affordable solutions when conducting large screening programs for early detection of ovarian cancer.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-ToF MS) is a very sensitive, affordable, and accurate technique for mass determination of biomolecules [10–13]. It enables precise detection of ionized peptides and proteins represented as peaks of mass-to-charge ratio (m/z). This technology has been considered to have high potential for clinical diagnostics and is already used in clinical microbiology [10,13]. One of the main reasons for this is the reduced cost per test in comparison with genetic-based testing and most currently used microbiological techniques for clinical identification of infection [11,12,14]. Such 10-fold reductions in cost can enable screening programs to be rolled out, even in poorer countries. In the last years, MALDI-ToF MS has further evolved to become an ultra-fast high-throughput technology suitable for diagnostic screening purposes, with the capacity of each machine to generate approximately seven mass spectra per-patient samples per minute [15]. Thus, this technology would be ideal for mass screening at-risk ovarian cancer populations within a sensible and reasonable processing time [11,12,14]. Applying MALDI-ToF technology to ovarian cancer detection has previously been attempted using blood serum samples with an optimal sensitivity of 71% and specificity of 68% [16]. However, although promising, this is still far from an ideal solution for population screening purposes.

More recently, MALDI-ToF application for human disease diagnostics has been enhanced by cutting-edge automated bioinformatics pipelines for spectral analysis [17–20]. This approach relies on the detection of mass spectral pattern recognition rather than specific biomarkers, which is the equivalent of simultaneously detecting multiple relevant biomarkers. Using mathematical models for disease scoring and machine learning approaches, these mass spectral patterns have been successfully applied to the detection of aneuploidies in spent IVF-embryo/blastocyst media and fetal Down's syndrome from the urine of pregnant women, with high sensitivity and reasonable specificities [20–22]. Further, these pipelines have been developed enabling automated and ultra-fast mass spectral data processing within seconds, creating software tools that enable screening for diseases in general populations by clinical laboratories [18,19]. Thus, it is highly feasible that this MALDI-ToF-based technology can be deployed for early-stage detection of ovarian cancer.

This study set out to test the hypothesis that this technological approach can provide a more suitable solution for general population screening programs. To address this question,

we conducted a retrospective cohort study using blood serum samples from ovarian cancer patients in the development of MALDI-ToF MS-based diagnostic tools, for early- and late-stage cancer detection. Here, we report the diagnostic power and performance of these tools using distinct spectral profiles of MALDI-ToF MS generated from two experimental approaches, optimized using advanced bioinformatics and machine learning.

2. Materials and Methods

2.1. Study Samples and Patients

A total of 181 anonymized human blood serum samples (100 μ L) were obtained from a commercial stored biological sample biobank (Invent diagnostica, Berlin, Germany).

No personal identifiers were supplied, only sample numbers with metadata as to ethnicity (if known), clinical disease stage, and age at the time the sample was taken. Of these serum samples, 38 were collected from healthy age-matched women, which were set as the control cohort of the study. The remaining 142 were collected from women who were diagnosed with ovarian cancer at a particular clinical stage of disease: 37 samples were grouped into ovarian cancer early-stage disease (FIGO) 1 and 2, 39 samples were stage 3, and 67 samples were stage 4 group. The ovarian cancer diagnostics were confirmed by the biobank as clinical staging but confirmed by histopathology where possible/available. All serum samples were shipped to MAP Sciences (Bedford, UK) for further laboratory processing and mass spectral analysis. The study was approved by the Internal Research Board of MAP Sciences and independently reviewed by the Ethics Committee of NISAD (Lund) Sweden.

2.2. Low Versus High Mass Spectral Pattern Analysis

MALDI-ToF mass spectrometry is versatile in that it can resolve small organic molecules and peptides (i.e., predominately metabolites) up to 15–20 kD, as well as large protein glycoproteins and complexes up to 200 kD or more. However, this requires optimization in both process chemistry/matrix selection and instrument settings in order to generate reliable and reproducible qualitative and (semi-)quantitative data in the different mass ranges.

2.2.1. Low Mass (2000 to 20,000 m/z) Spectra Range Data Generation

Mass spectra were generated using α -Cyano-4-hydroxycinnamic acid (CHCA) matrix. The MALDI-8020 (Shimadzu–Kratos, Manchester, UK), MALDI-ToF mass spectrometry instrument was externally calibrated using Bradykinin Fragment 1–7 (756.4 Da) (ProteoMass™, Supelco, Sigma-Aldrich, Bellefonte, PA, USA). Samples were diluted to 1:80 with mass-spectrometry-grade water (Romil, Cambridgeshire, UK) and 1 μ L plated on a 48 well stainless-steel target plate using a sandwich technique [23]. Mass spectral data were generated in positive linear mode with a 100 Hz solid-state laser (355 nm) at a mass range of 200 to 2000 m/z ; blanking at 150 m/z , and pulse extraction set at 700. A square raster pattern consisting of 500 position and 17 penetration shots was employed to sample across the entire spot. Thus, 8500 spectra per sample were collected in order to achieve a statistically consistent average spectra per patient sample. Averaged data were exported in mzML format. For each patient's sample, we generated three replicates and compared the obtained averaged mass spectra to ensure reproducibility and high-quality assessment.

2.2.2. High Mass (20,000 to 200,000 m/z) Spectra Range Data Generation

Mass spectra were generated using sinapinic acid (SA) matrix. The MALDI-8020 (Shimadzu–Kratos, Manchester, UK), MALDI-ToF mass spectrometry instrument was externally calibrated using Apomyoglobin (16,952 m/z) (ProteoMass™, Sigma-Aldrich). Samples were diluted to 1:80 with mass-spectrometry-grade water (Romil, Cambridgeshire, UK) and 1 μ L plated on a 48 well stainless-steel target plate using a sandwich technique [23]. Mass spectral data were generated in positive linear mode with a 100 Hz solid-state laser (355 nm) at a mass range of 20,000 to 200,000 m/z , blanking at 19,000 m/z , and pulse extraction set at 23,000. A square raster pattern consisting of 500 position and

17 penetrations shots was employed to sample across the entire spot. Thus, 8500 spectra per sample were collected in order to achieve a statistically consistent average spectra per patient sample. Averaged data were exported in mzML format. For each patient's sample, we generated three replicates and compared the obtained averaged mass spectra to ensure reproducibility and high-quality assessment.

2.3. Data Processing

Mass spectral data generated by MALDI-ToF were systematically processed using an automated computational pipeline based on the previously developed workflow [18,22]. This pipeline was modified from the previous workflow [18,22] and optimized for serum samples prepared with CHCA and SA matrices. Using this workflow, the mass spectra were corrected for the baseline and normalized to total ion count as described previously [18]. The pipeline was modified to extract well-defined peak locations in the m/z region and calculate their relative heights as proxies for signal intensity. This was achieved by inserting in the pipeline a variable m/z window scan to identify optimal threshold parameters for peak resolution detection. The peak resolution was defined by the degree of proximation of the adjacent valleys to the baseline by using a maximum value identification within the peak valleys, thus enabling identification of peaks within variable ranges of m/z windows and ignoring peaks within shoulders. To ensure the quality of peak identification, mass spectral quality degree was computed based on the average signal-to-noise ratio estimated using the computed local noise on each mass spectrum [18,19,24]. Mass spectral quality decision (good/bad) was inferred by comparing a threshold value that was obtained by carefully observing two sets of mass spectra generated under the same conditions. A minimum threshold of 10 well-defined peaks was implemented for quality control decisions to discard data without enough peaks for pattern detection. A function for generating the average mass spectrum was added to the pipeline and the percentual variation between averages computed.

2.3.1. Comparative Analysis and Scoring

Peak heights detected within a particular m/z region were systematically compared between each stage and the control group for a variable range of m/z windows from 5 to 500 m/z . Statistical significance between stages and control peaks was tested using the non-parametric, two-tailed Wilcoxon–Mann–Whitney test. Median signal intensities (MIs) were computed for each m/z region and were compared with the control if the overlap between these groups was lower than 15% as previously described [18]. The enrichment of peaks within a particular m/z region was also computed for each stage in comparison to the control group as previously described [18,20]. Peaks exclusively present in cancer stages and not in the control were also searched based on the computed enrichment values. Based on the above-mentioned comparative metrics, we developed a scoring-based mathematical model for ovarian cancer stage classification (OCS) with the following mathematical equation:

$$OCS = \sum_{i=1}^n rDIM_i \times E_i + \sum_{j=1}^m E_j$$

In this mathematical model $rDIM_i$ is the relative deviation between the median mass spectral signal intensity (peak height) characteristic of a particular stage on a particular m/z region i ; E_i is the peak enrichment on a particular m/z region i ; n is the total number of mass spectral intensity differences characteristic of a particular stage; E_j is the enrichment of a peak exclusively present in a particular stage; and m is the total number of characteristic peaks exclusively found in a stage.

2.3.2. Classification Model Extraction

The OCS mathematical equation was used for scoring and binary classification of ovarian cancer stages based on estimating optimal cutoff values. Multiple mass spectral signatures (peak presence and peak intensity differences) that compose the classification algorithm (OSC) were identified using the machine learning algorithm EvA-3 previously developed by the Bioenhancer Systems project. EvA-3 is a tree-based evolutionary hybrid algorithm that learns how to synergistically combine $n+1$ mass spectral features, ensuring gains on the pick-up rate and preventing an increase of the false positive with iteration time. A training set consisting of 50% of randomly selected data was separated. EvA-3 was programmed to learn mass spectral features on the training set for a maximum of 6 h of learning time and evolve the model for its maximum pick-up rate. Using EvA-3 the sensitivity, specificity, and the receiver operating characteristic (ROC) curve area under the curve (AUC) was computed using 100% of the data set.

2.3.3. Bioinformatics Pipelines

All data processing steps, optimizations, and predictive model developments were implemented into a single computational pipeline that systematically made the comparisons, generated all data visualizations, and generated the optimal predictive models. All data processing steps, together with ovarian cancer stage classification models, were also implemented into another pipeline. This final pipeline was developed for automated processing of multiple raw data mzML files and generating a binary classification prediction for each stage with respective scores on an output csv file. Both pipelines were implemented using code written in Python version 3.8. All data visualizations were generated using a matplotlib python library.

3. Results

Mass spectra of serum samples, from women with and without ovarian cancer, were generated using two protocols for MALDI-ToF MS for blood serum. For one, we used a CHCA matrix and for the other an SA matrix. Protocols were optimized for taking an average time of 80 minutes to generate a batch of 48 mass spectra (100 s per sample). We analyzed the mass spectral data generated from these protocols using our bioinformatics data pipeline. Most of the data generated through these experimental protocols were of high quality, 96% using CHCA matrix and 99% using SA matrix. The average mass spectra from patients with ovarian cancer at early and late stages were compared (Figure 1). These spectra are characterized by having more pronounced peaks located between 200 m/z and 700 m/z using CHCA matrix (Figure 1A) and between 2000 m/z and 5500 m/z using SA matrix (Figure 1B). Multiple spectral changes between cancer stages and controls can also be observed across the entire mass spectrum, including mass regions of lower intensities (box I and II in Figure 1). Proportionally, the averaged mass spectrum of early and late stages of ovarian cancer has substantial variation in comparison to women without cancer in particular mass regions across the entire mass spectrum (Figure 2). Some regions show distinct variations among stages, which may potentially be used for classification purposes. However, there is a substantial degree of complexity of these variations across the spectrum and overlap between stages. In addition, one order of magnitude of intensity variation was found across the data set. This further complicates a direct application of mass spectral variations for classification purposes.

To simplify the mass spectral data for developing classification models, we focus on analyzing well-defined peaks, which were used as signatures for the pattern-based scoring function for the likelihood of having a particular stage of ovarian cancer. First, we optimized the extraction of well-defined peaks on the data by increasing the number of detected peaks while reducing the variability of total peaks detected (Figure 3). We obtained an optimal increase in well-defined peak distribution ranging from 26 to 57 detected peaks in each sample for CHCA matrix mass spectral data (Figure 3A). For SA matrix mass spectral data, the optimization reduced the number of peaks for a final distribution ranging

from 12 to 113 detected peaks in each sample (Figure 3B). Secondly, on this resulting data, we applied machine learning for training and validating predictive models based on mass spectral patterns, to optimize the binary classification of ovarian cancer detection on early stages (1 or 2) and late stages (3 and 4). We obtained six classification models, one for each type of matrix explored and specifically for detecting ovarian cancer at early stages (stage 1 or 2), stage 3, and stage 4. The performance evolution of these models is presented in Figure 4. Our machine learning approach enables the evolution of predictive models with good to excellent classification power ($98\% < \text{AUC} > 70\%$) for ovarian cancer stage detection [25,26]. The maximum obtained sensitivities of these models were between 92% and 100% using CHCA and 100% for SA matrix. This indicates high potential for detecting ovarian cancer in early and late stages. However, models generated using SA matrix data show a much higher specificity compared with the ones obtained using CHCA matrix, having a much lower false positive rate. Among SA-generated models, we obtained final classification specificities of 92% for early stages, 96% for stage 3, and 82% for stage 4.

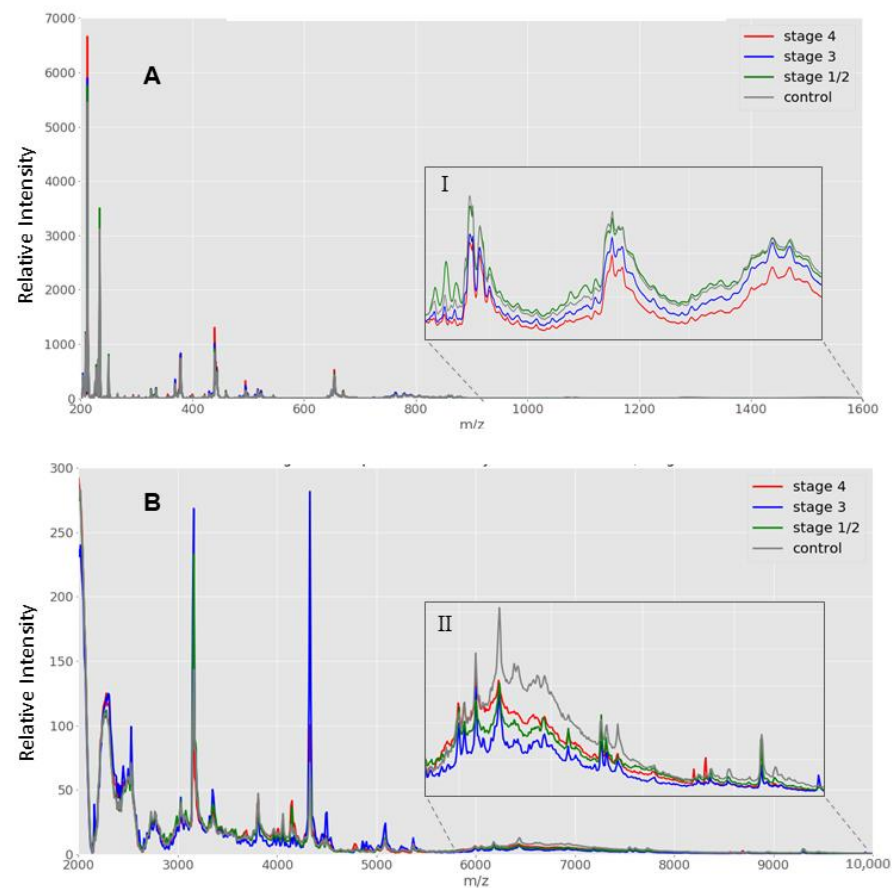


Figure 1. Averaged mass spectra of serum samples from women diagnosed with ovarian cancer at different stages and controls. (A) Spectrum generated with CHCA matrix. (B) Spectrum generated with SA matrix. Red indicates averaged mass spectra of women with ovarian cancer at stage 4, blue at stage 3, and green at early stages (stage 1 and 2). Gray spectrum indicates averaged mass spectrum of women with no ovarian cancer. Relative intensities are expressed in ppm. Lower-intensity region spectrum zooms from regions within dashed lines are depicted in the inset spectra I and II for CHCA and SA matrix, respectively.

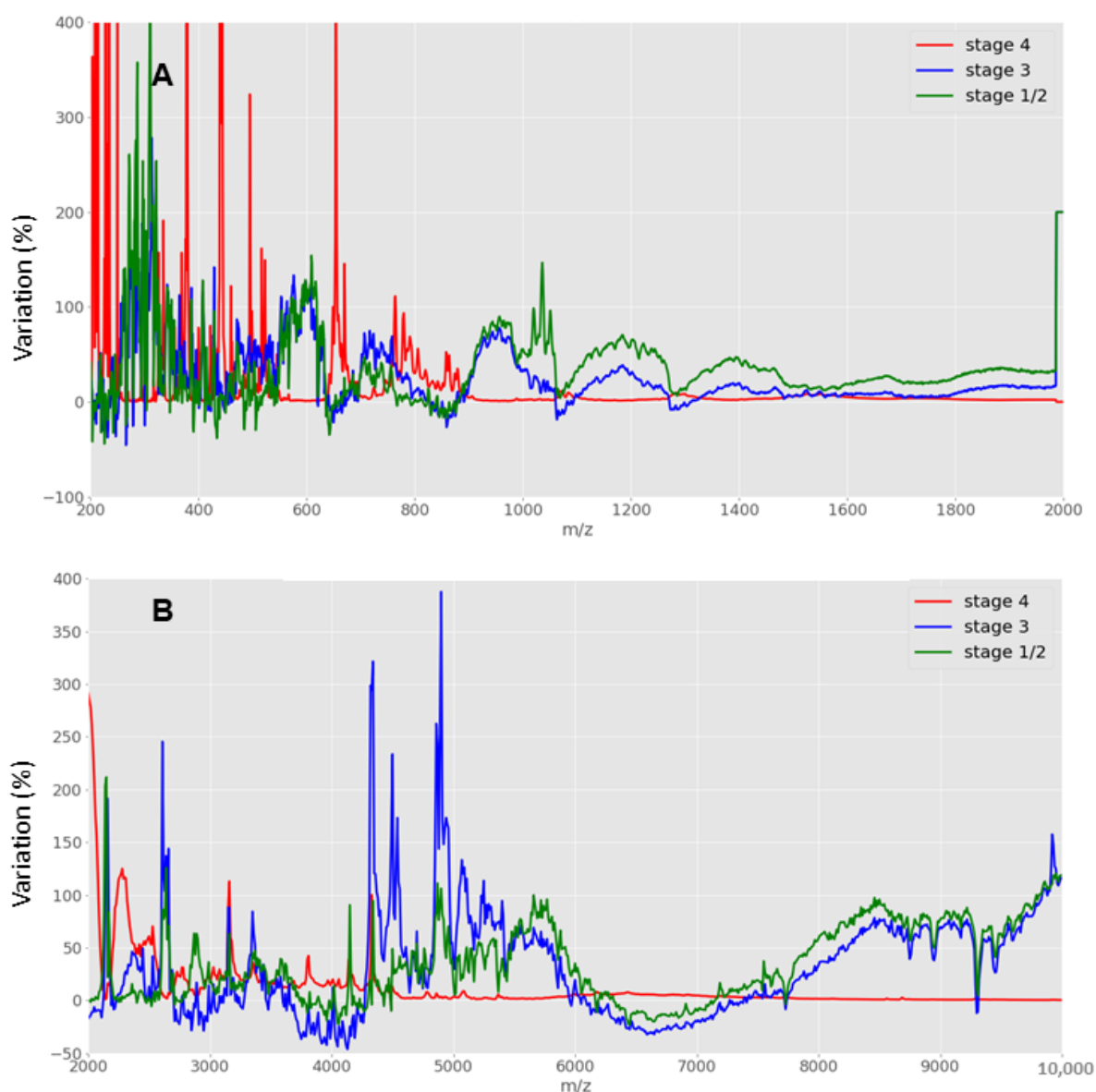


Figure 2. Averaged variation across the mass spectrum of women with ovarian cancer in comparison with the control group. (A) Variation calculated for CHCA-matrix-generated data. (B) Variation calculated for SA-matrix-generated data. Red indicates variations in women with ovarian cancer at stage 4, blue at stage 3, and green at early stages (stage 1 or 2). Variations were calculated by the difference between the averaged mass spectrum of a particular stage and the control (not having cancer), expressed in percentage of the control.

The selected spectral signatures learned during model evolution enabled us to obtain mass spectral patterns of the ovarian cancer stages that have a high classification potential. The patterns are composed of complex combinations of multiple characteristic peaks detected together, with multiple intensity variations (increase or decrease) in discrete mass spectral ranges (Figure 5). Patterns differ among stages with only few signatures that have similar qualitative variation. Analysis of the individual contribution of each learned signature for the overall model specificity enables us to pinpoint the ones that impact specificity the most (Figure 5). This information is useful for further tuning of the predictive models to balance the sensitivity and specificity of final deployed models. Thus, we further refined the models by neglecting the last learned pattern components, which increased the false positive rates, rendering a final model version with 97% specificity for SA-derived models and 86% specificity for CHCA-derived models (see Figure 4).

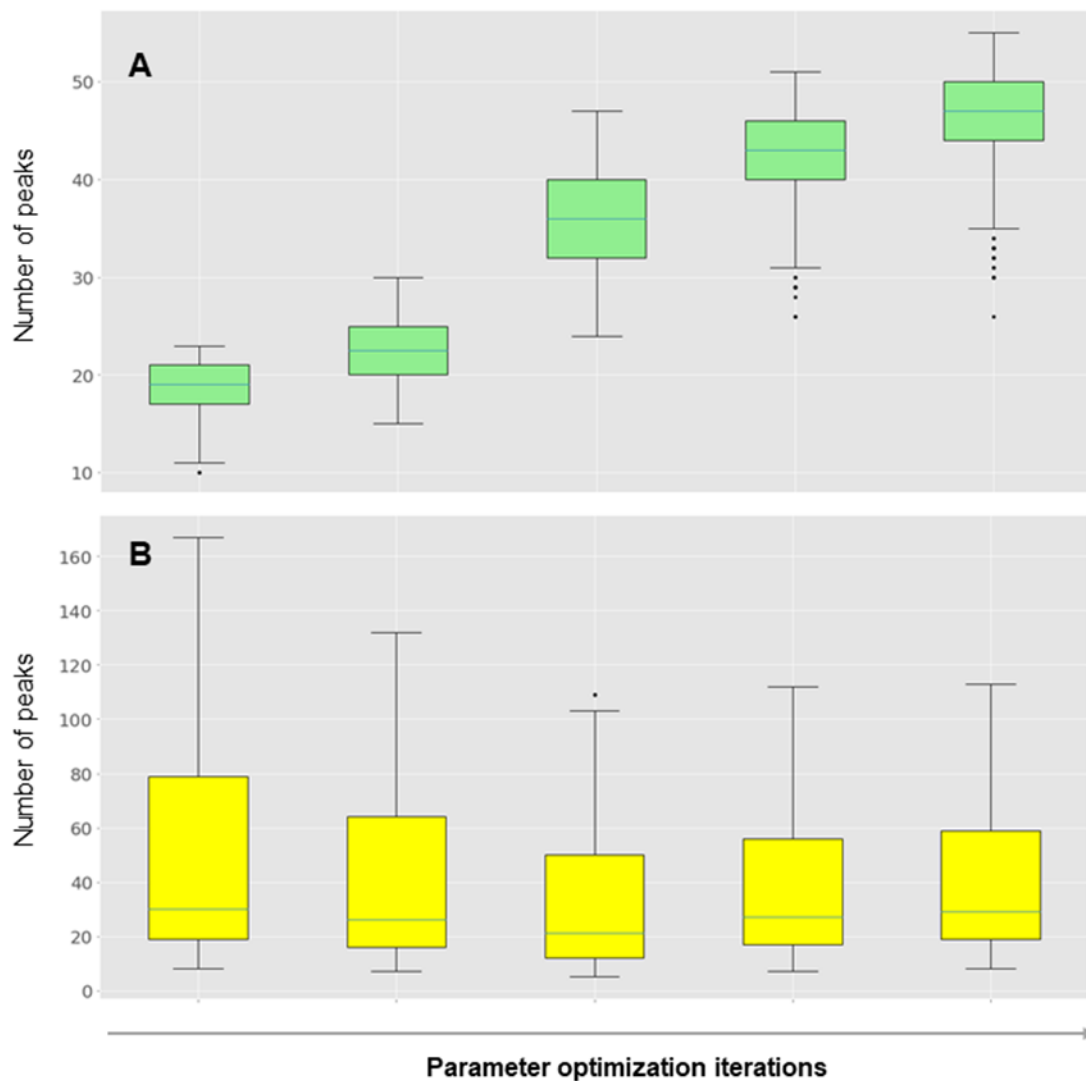


Figure 3. Bioinformatic pipeline optimization of well-defined peak detection for ovarian cancer classification. **(A)** CHCA-matrix-generated data optimization. **(B)** SA-matrix-generated data optimization. Optimization was conducted by systematic combination of two peak-detection parameters, the percentage of tolerance of well-defined peak (TWP), and minimum distance between peaks for peak finding (MD). The ranges for scanning for TWP were set from 95% to 85%. For the MD, the scanning ranges were set from 20 m/z to 50 m/z for CHCA data and from 300 m/z to 500 m/z for SA data. Parameters were systematically combined with an algorithm to either fetch gains on maximum peaks detected or reduce variations between samples.

Finally, we implemented the final refined version of the classification models into prototype pipelines, one for scoring cancer stages on mass spectral data generated using CHCA matrix and another for scoring SA matrix. Running these pipelines on the entire data set, we obtained an average data processing time to score all stages of 1.4 seconds per sample for CHCA data and 29.7 seconds for SA data. Using the selected cutoffs obtained during model evolution for refined versions, we have recapitulated the same model performance shown in Figure 4 as expected (see Supplementary Tables S1 and S2). The obtained score's distribution was compared (Figure 6) showing higher scoring differences among stages and the control for SA-matrix-derived models in comparison with CHCA-matrix-derived models. The results further show that models designed for detecting a particular stage often score higher than cutoffs on other stages in comparison with controls (see the example of SA-derived models in Figure 7). This generates multiple hits for predicting ovarian

cancer at different stages, indicating that our scoring models have low stage specificity. Yet, best scoring discrimination between a stage and the control is observed for the model that was optimized for that stage. Interestingly, the combination of algorithm predictions using a simple decision rule of at least one positive hit for cancer, regardless of stage, has improved the sensitivity of the individual algorithm to 94% for CHCA and 99% for SA (Supplementary Tables S1 and S2, respectively). However, the rule base combination of algorithms also reduced the specificity to 92% for SA- and 71% for CHCA-derived models.

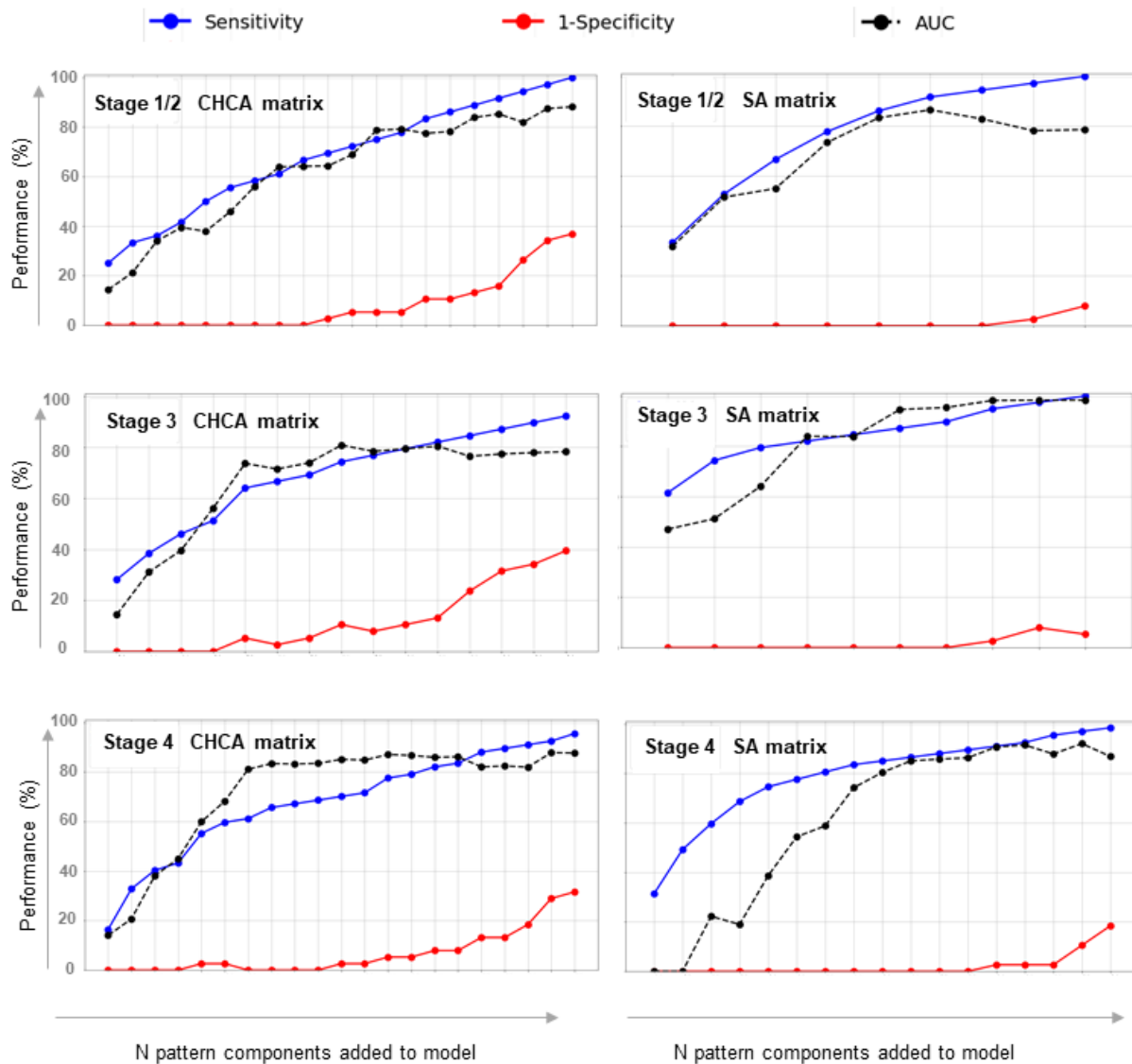


Figure 4. Performance evolution of predictive models for ovarian cancer stage classification. Each panel corresponds to an algorithm learning mass spectral signatures for a particular cancer stage using a given matrix type (indicated within a panel). Identification of peak signature combinations was performed by the EvA-3 algorithm, which selectively adds a mass spectral feature, updates the performance, and chooses the next best combination until further performance optimization is no longer possible. The algorithm starts with the best detected single signature (N) and each subsequent point is an N+1 combination of signatures. Dots indicate the computed values for each combination and respective performance indicators. Sensitivity evolution is depicted in blue lines, specificity in red lines, and the area under the curve of the computed ROC (AUC) in black dashed lines.

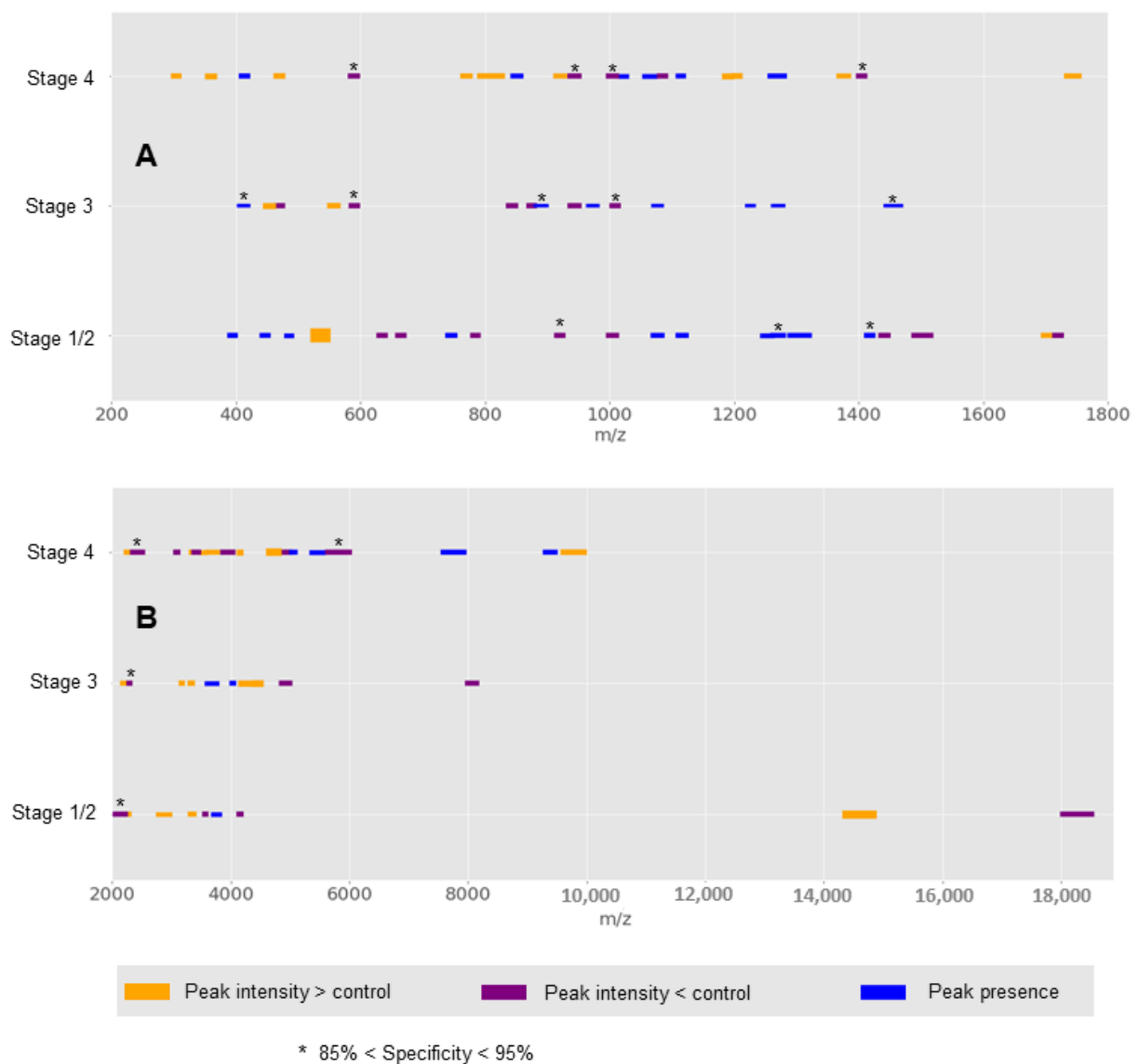


Figure 5. Mass spectral patterns of ovarian cancer classification models. (A) Patterns obtained for CHCA matrix mass-spectral-data-derived models. (B) Patterns obtained for SA matrix mass-spectral-data-derived models. Classification models correspond to final models evolved through the machine learning process. Orange indicates significant increase in mass spectral intensity in comparison with the control group. Purple indicates a significant reduction in mass spectral intensity in comparison with the control group. Median intensity variation degrees are shown by the broadness of the m/z regions, scaled to be visible for 1 order of magnitude variations. Blue indicates peaks that are characteristic for a given stage (peak presence). The * indicates the mass spectral regions that have an impact on the model's specificity.

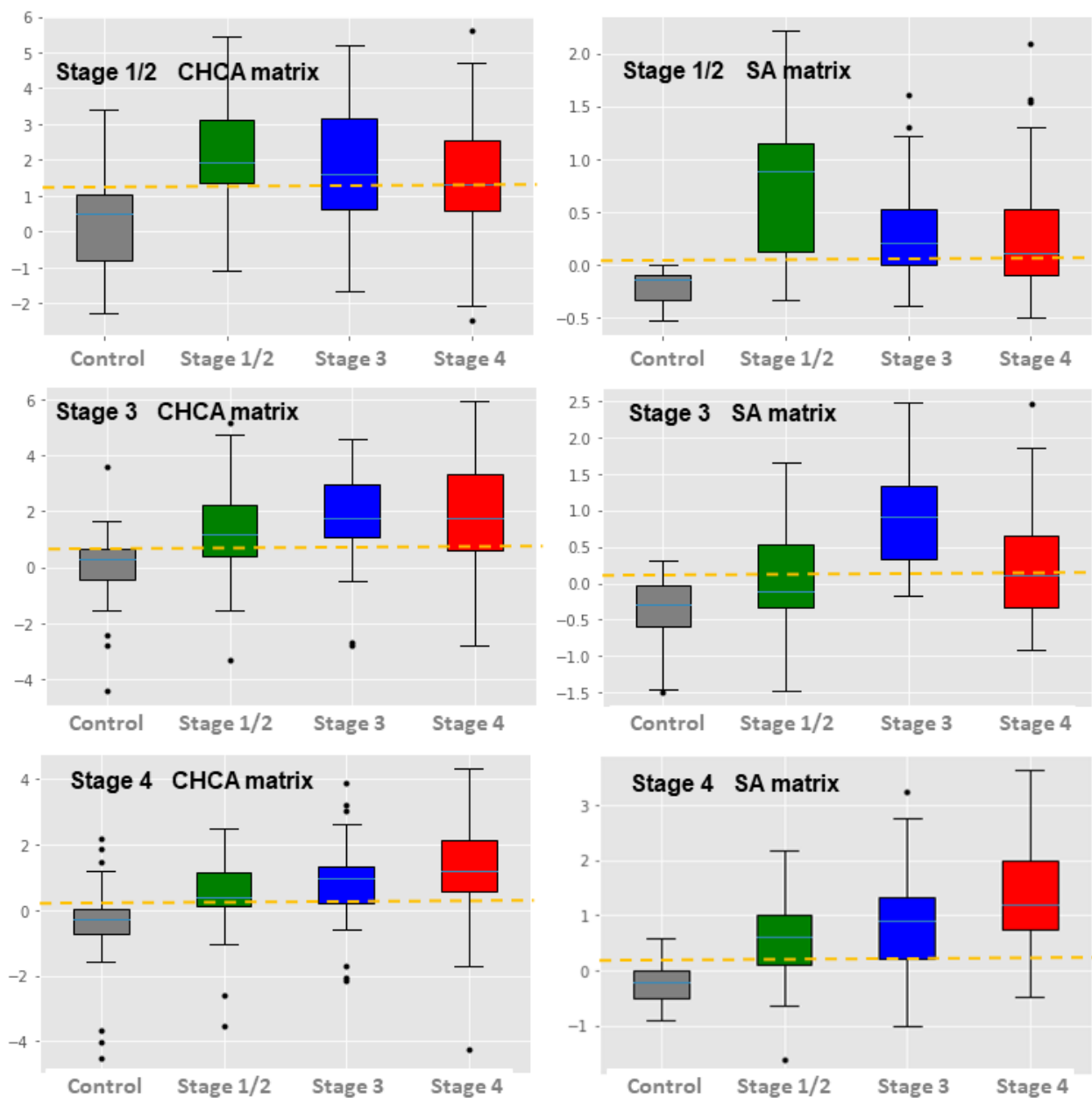


Figure 6. Stage-specific scores of serum mass spectra using each algorithm. Each panel corresponds to the scores using a classification model specifically designed for a particular cancer stage and using a given matrix type (indicated within a panel). Each box represents scores of mass spectral data of distinct groups (Control, Stage 1/2, Stage 3, and Stage 4) and, respectively, depicted with different colors (grey, green, blue, and red). The classification cutoffs obtained through machine learning for each model are depicted by the orange dashed lines. Medians are indicated as blue lines; whiskers represent 95% of distribution and black dots the outliers.

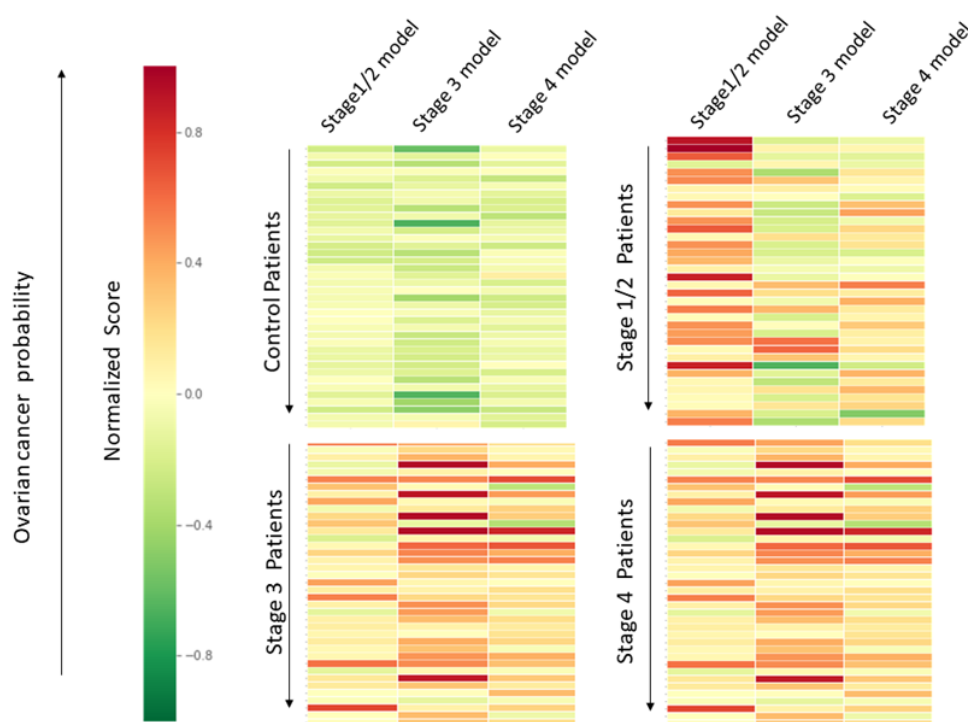


Figure 7. Comparison of the predictive scores of controls and those diagnosed with distinct cancer stages using the SA-based stage-specific models. Scores were normalized to reflect the deviation from the cutoff value within the maximum score values detected (maxScore) with the mathematical expression $(\text{patient score} - \text{cutoff}) / \text{maxScore}$. Value of zero indicates that patients' scores are equal to the model cutoff, positive values above cutoff, and negative values below cutoff. Values 1 and -1 indicate that patient scores are the maximum score observable. Scores of patients and controls are shown as grouped heatmaps and the colors indicate the normalized scores. Red indicates the highest score and green the lowest score for predicting cancer using a stage-based model. The correlated probability of cancer based on scoring is indicated as an arrow next to the color gradient legend. Low probability starts from dark green and ends in high probability at dark red.

4. Discussion

Failure to detect ovarian cancer at an early stage has motivated many researchers to attempt improving current diagnostic tests or develop novel ones [2–4]. In this study, we developed and optimized a MALDI-ToF mass spectral pattern-based bioinformatics pipeline, which enabled the detection of ovarian cancer in blood serum with high performance. In comparison with the reported sensitivities for the CA-125 test, used worldwide in screening programs, our approach provides a substantial gain in sensitivity and specificity for early- and late-stage disease detection [3,5]. These gains were almost of 50% in sensitivity for early-stage detection, suggesting that implementing our MALDI-ToF approach in screening programs would have a huge impact in preventing thousands of misdetected cases every year [2,3].

Other recently developed solutions to the problem are based on longitudinal biomarker level monitoring, statistical models with additional bio-markers and microRNA quantification. In particular, measurement of plasma circBNC2 has enhanced the performance of CA-125 surveillance/monitoring boosting ovarian cancer detection [5,8,9]. For early-stage disease, these approaches reported optimal sensitivities ranging from 82% to 92% with specificities between 86% and 95%. For later-stage ovarian cancers, longitudinal models and circBNC2-based tests report sensitivities of 90% to 100% with specificities of 84% to 95%. In comparison, our MALDI-ToF-based solution demonstrated here has higher sensitivity (95% to 99%) with equally higher specificities (97%), representing a significant gain. Our best solution outcompetes the performance of the recent approach of combin-

ing multiple longitudinal risk models which, although increasing sensitivity, report only equivalent specificity for early-stage ovarian cancer [9].

This was possible due to the fine tuning of the MALDI-ToF MS data pattern recognition models towards improving the specificity, whilst keeping high sensitivity by removing particular signatures in the patterns that contributed to an increase in the number of false positives (Figures 4 and 5). Thus, our MALDI-ToF-based approach provides an optimal solution and framework for maximizing the gain of true positives while minimizing the false positives. Balancing these is useful for designing optimal ovarian cancer screening programs to prevent unnecessary, invasive investigations in false positive cases or delayed diagnosis in the cases of false negatives [2,4,5].

Interestingly, the models developed for a given stage have also scored positive for other stages and the combination of predicted outcomes using a decision rule further improved the cancer detection performance. This could be due to multiple intermediate cancer stages that share common mass spectral profiling or a continuous mass spectral profiling evolution that correlates with cancer development. Further research is still necessary for understanding how to correlate mass spectral profiling with cancer development to enable better stage classification or establishing a correlation with cancer aggressiveness.

The application of MALDI-ToF mass spectrometry for ovarian cancer detection on blood serum samples has previously been attempted, but the results have been poor and with experimental bias [16,27]. In a recent systemic review and meta-analysis of 18 studies evaluating the accuracy of MALDI-ToF MS for ovarian cancer, the reported overall sensitivity was 77% (95% CI: 73–80%) and specificity was 72% (95% CI: 70–74%) [28]. In a more recent study, Swiatly et al. followed a traditional proteomic approach for multiple biomarker detection using high-performance liquid chromatography (HPLC) applying conventional statistical models [6,16,29,30]. This study reported only a 71% sensitivity with a 68% specificity. For classification purposes, we followed quite a different approach. Instead, we applied mass spectral pattern recognition and used novel machine learning algorithms to optimize predictive models [17,31,32]. The substantial improvement in performance demonstrates a methodological breakthrough, which launches MALDI-ToF mass spectrometry as a highly sensitive technique for ovarian cancer diagnostics.

The reduced cost of running MALDI-ToF-based technology is a huge advantage for the implementation of our proposed solution for ovarian cancer screening programs. This advantage is mainly due to the reduced costs of reagents and the possibility of running multiple samples in one single run [11,14,19]. In our approach we have not used HPLC for biomarker resolution prior to mass signature detection in a mass analyzer, as is traditionally employed in most mass spectral techniques in proteomics and metabolomics [6,29,30]. Thus, this ‘dilute and shoot’ direct mass analysis enables further economization of operational costs, which are otherwise considerable from reagents and additional equipment. Indeed, the total estimated costs per test using our approach would be five to ten GBP, given a sample throughput of 50,000 a year over three years. It would approximately halve the cost when compared to the CA-125 screening test [7,33]. According to recent analysis, the major benefit to health care systems for each person’s case of early versus late stage detection of ovarian cancer is the saving of approximately 90 thousand GBP per year [33]. Thus, the impact of implementing MALDI-ToF technology with our approach for screening would definitely improve current screening programs in terms of cost-effectiveness and long-term health economics. This is not the case for longitudinal follow-up monitoring models with additional biomarkers or circBNC2-based tests, since these techniques require the purchase of more expensive reagents [11,14].

However, there is still the need for governmental or private investment for implementing MALDI-ToF technology into clinical laboratories as each piece of equipment costs hundreds of thousands of GBP [11,14]. However, it can be argued that this investment is absolutely necessary, and the savings will pay off in the short-term, given that the hardware of this technology is used in other diagnostics such as microbiology and is being applied in the diagnosis of multiple other diseases [14]. Future implementation of MALDI-ToF-based

technology would enable government support for a broad number of screening programs, in particular in poorer countries [4,5,14].

Another advantage of MALDI-ToF mass spectrometry for disease screening purposes is its capacity for generating multiple spectra in the scale of minutes [17,19]. We have further optimized and automatized the testing using bioinformatics, which rendered multiple results at a rate of a few seconds per sample. Excluding the time of sample preparation and calibration, each MALDI-ToF mass spectrometer has an estimated maximum capacity of generating approximately 620 results per day, with a potential increase if optimizing for speed of data acquisition. At this rate, only 102 MALDI-ToF machines working full-time would be required to screen the entire UK female population over 30 in one year [34]. Thus, the approach could make it feasible to screen an entire relevant population group with an initial investment for laboratory setup or hiring specialized companies that already have MALDI-ToF technology.

In our work, we also optimized, tested, and compared the potential application of CHCA and SA matrices for ovarian cancer diagnostics. These have not been compared before for this purpose [16]. Our results have pinpointed better diagnostic potential for SA-derived models in comparison to the CHCA. This was unexpected as CHCA matrix is known to generate more robust spectra in comparison to SA in urine, blood, serum, and culture media [18,20–22,35]. Comparatively, mass spectral data acquired using SA matrix result in richer data; however, the acquisition simultaneously produces more background noise, often leading to the complex baseline that requires more sophisticated corrections [18]. This has previously been solved using automated bioinformatics pipelines and is successfully applied now for ovarian cancer mass spectral analysis [18].

There are several limitations to our study. First, our predictive models were not capable of providing an accurate classification between stages. This was evident during testing with multiple models on all data regardless of the diagnosed stage. Nevertheless, the low intra-stage specificity observed is irrelevant for screening purposes as other confirmatory methods such as TVUS offer support to further clinical diagnostics and therapy assessment [2,4,5]. These can be applied as a second-stage evaluation following positive scoring by inexpensive serum MALDI-ToF MS screening [2,4]. Secondly, huge deviations in predictions can be obtained if experimental protocols, matrices, and mass spectrometer settings are not exactly followed, including sample preparation and storage. This has been observed before for other studies of data modeling using both CHCA and SA matrices [18–21,36]. The reason for this is the high precision needed for our pattern recognition. Peak widening on a mass spectrum with experimental conditions and calibrations is detrimental [18,32,37]. Third, the number of volunteers in different cancer stages and the control were limited for conducting more robust validations of predictive models [26]. As the numbers for stages 1 and 2 were reduced, we had to aggregate these stages into one to have statistical meaning. In an ideal scenario, the numbers of volunteers within each group should be approximately a thousand, in order to split into large training and validation subsets [26]. Consequently, the reported performance may slightly change if more data were to be tested.

5. Conclusions

In this study, we developed a novel MALDI-ToF-based approach for ovarian cancer detection in serum samples. Using our state-of-the-art bioinformatics and machine learning, we have successfully optimized MALDI-ToF technology for fast, fully automated, and highly sensitive ovarian cancer detection. This was demonstrated in the detection of both early- and late-stage ovarian tumors. This is in stark comparison with currently available technologies, which lose performance significantly when screening for early-stage disease.

Importantly, we demonstrated that MALDI-ToF technology can potentially be applied for the screening of entire populations with acceptable costs and timings. This is the methodological breakthrough that can be implemented as a future diagnostic test. Further clinical screening studies are required to confirm the efficiency of ovarian cancer detection, leading to implementation of robust MALDI-ToF-based technology for ovarian

cancer. Nevertheless, with increasing evidence of the MS assay's diagnostic accuracy and instrument robustness, it has become imminent to consider MS in the clinical application for cancer screening.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app12063030/s1>, Table S1: CHCA-derived stage-specific model predictions and rule-based cancer predictions, Table S2: SA-derived stage-specific model predictions and rule-based cancer predictions.

Author Contributions: Conceptualization, R.J.P., R.Z. and R.K.I.; methodology, R.J.P., R.Z., J.C.L., C.S.J. and R.K.I.; formal analysis, R.J.P., R.Z., J.C.L., C.S.J. and R.K.I.; investigation, R.J.P., R.Z., J.C.L., C.S.J. and R.K.I.; writing—original draft preparation, R.J.P.; writing—review and editing, R.Z., J.C.L. and R.K.I.; visualization, R.J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Internal Research Board of MAP Sciences and independently reviewed by the Ethics Committee of NISAD (Lund) Sweden.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Restrictions apply to the availability of these data.

Conflicts of Interest: R.J.P. declares no conflict of interest. R.Z., C.S.J. and J.C.L. are employees of MAP Sciences Ltd. R.K.I. has filed patents on MALDI ToF mass spectral profiling. R.K.I. declares a potential conflict of interest through part-ownership of shares in MAP Sciences Ltd.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [[CrossRef](#)] [[PubMed](#)]
2. Aust, S.; Seebacher-Shariat, V. Screening for ovarian cancer: Is there still hope? *Mag. Eur. Med. Oncol.* **2020**, *13*, 189–192. [[CrossRef](#)]
3. Rauh-Hain, J.A.; Krivak, T.C.; del Carmen, M.G.; Olawaiye, A.B. Ovarian cancer screening and early detection in the general population. *Rev. Obstet. Gynecol.* **2011**, *4*, 15–21. [[CrossRef](#)] [[PubMed](#)]
4. Patni, R. Screening for ovarian cancer: An update. *J. Midlife. Health* **2019**, *10*, 3–5. [[CrossRef](#)]
5. Menon, U.; Griffin, M.; Gentry-Maharaj, A. Ovarian cancer screening—Current status, future directions. *Gynecol. Oncol.* **2014**, *132*, 490–495. [[CrossRef](#)]
6. Kong, A.A.; Gupta, C.; Ferrari, M.; Agostini, M.; Bedin, C.; Bouamrani, A.; Tasciotti, E.; Azencott, R. Biomarker Signature Discovery from Mass Spectrometry Data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2014**, *11*, 766–772. [[CrossRef](#)]
7. Gorski, J.W.; Quattrone, M.; van Nagell, J.R.; Pavlik, E.J. Assessing the Costs of Screening for Ovarian Cancer in the United States: An Evolving Analysis. *Diagnostics* **2020**, *10*, 67. [[CrossRef](#)]
8. Hu, Y.; Zhu, Y.; Zhang, W.; Lang, J.; Ning, L. Utility Of Plasma circBNC2 as a Diagnostic Biomarker In Epithelial Ovarian Cancer. *Onco. Targets. Ther.* **2019**, *12*, 9715–9723. [[CrossRef](#)]
9. Whitwell, H.J.; Worthington, J.; Blyuss, O.; Gentry-Maharaj, A.; Ryan, A.; Gunu, R.; Kalsi, J.; Menon, U.; Jacobs, I.; Zaikin, A.; et al. Improved early detection of ovarian cancer using longitudinal multimarker models. *Br. J. Cancer* **2020**, *122*, 847–856. [[CrossRef](#)]
10. Angeletti, S. Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. *J. Microbiol. Methods* **2017**, *138*, 20–29. [[CrossRef](#)]
11. Gaillot, O.; Blondiaux, N.; Loiez, C.; Wallet, F.; Lemaître, N.; Herwegh, S.; Courcol, R.J. Cost-effectiveness of switch to matrix-assisted laser desorption ionization-time of flight mass spectrometry for routine bacterial identification. *J. Clin. Microbiol.* **2011**, *49*, 4412. [[CrossRef](#)] [[PubMed](#)]
12. Tan, K.E.; Ellis, B.C.; Lee, R.; Stamper, P.D.; Zhang, S.X.; Carroll, K.C. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: A bench-by-bench study for assessing the impact on ti. *J. Clin. Microbiol.* **2012**, *50*, 3301–3308. [[CrossRef](#)] [[PubMed](#)]
13. Duncan, M.W.; Nedelkov, D.; Walsh, R.; Hattan, S.J. Applications of MALDI Mass Spectrometry in Clinical Chemistry. *Clin. Chem.* **2015**, *62*, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Iles, R.; Naase, M. Future Laboratory Medicine: Rapid, Efficient and Affordable Screening for Haemoglobinopathies by MALDI-ToF Mass Spectrometry. *Adv. Biochem. Biotechnol.* **2018**, *2018*, 1–5.
15. Haslam, C.; Hellicar, J.; Dunn, A.; Fuetterer, A.; Hardy, N.; Marshall, P.; Paape, R.; Pemberton, M.; Resemannand, A.; Leveridge, M. The Evolution of MALDI-TOF Mass Spectrometry toward Ultra-High-Throughput Screening: 1536-Well Format and Beyond. *J. Biomol. Screen.* **2016**, *21*, 176–186. [[CrossRef](#)] [[PubMed](#)]

16. Swiatly, A.; Horala, A.; Hajduk, J.; Matysiak, J.; Nowak-Markwitz, E.; Kokot, Z.J. MALDI-TOF-MS analysis in discovery and identification of serum proteomic patterns of ovarian cancer. *BMC Cancer* **2017**, *17*, 1–9. [[CrossRef](#)]
17. Pais, R.J.; Iles, R.K.; Zmuidinaite, R. MALDI-ToF Mass Spectra Phenomic Analysis for Human Disease Diagnosis Enabled by Cutting-Edge Data Processing Pipelines and Bioinformatics Tools. *Curr. Med. Chem.* **2020**, *28*, 6532–6547. [[CrossRef](#)]
18. Pais, R.J.; Zmuidinaite, R.; Butler, S.A.; Iles, R.K. An automated workflow for MALDI-ToF mass spectra pattern identification on large data sets: An application to detect aneuploidies from pregnancy urine. *Inform. Med. Unlocked* **2019**, *16*, 100194. [[CrossRef](#)]
19. Pais, R.J.; Jardine, C.; Zmuidinaite, R.; Lacey, J.; Butler, S. Rapid, Affordable and Efficient Screening of Multiple Blood Abnormalities Made Possible Using an Automated Tool for MALDI-ToF Spectrometry Analysis. *Appl. Sci.* **2019**, *9*, 4999. [[CrossRef](#)]
20. Pais, R.J.; Sharara, F.; Zmuidinaite, R.; Butler, S.; Keshavarz, S.; Iles, R. Bioinformatic identification of euploid and aneuploid embryo secretome signatures in IVF culture media based on MALDI-ToF mass spectrometry. *J. Assist. Reprod. Genet.* **2020**, *37*, 2189–2198. [[CrossRef](#)]
21. Iles, R.K.; Nicolaides, K.; Pais, R.; Zmuidinaite, R.; Keshavarz, S.; Poon, L.; Butler, S. The importance of gestational age in first trimester, maternal urine MALDI-ToF MS screening tests for Down Syndrome. *Ann. Proteomics Bioinform.* **2019**, *3*, 010–017. [[CrossRef](#)]
22. Sharara, F.; Butler, S.A.; Pais, R.J.; Zmuidinaite, R.; Keshavarz, S.; Iles, R.K. BESST, a Non-Invasive Computational Tool for Embryo Selection Using Mass Spectral Profiling of Embryo Culture Media. *EMJ Reprod. Health* **2019**, *5*, 59–60.
23. Dai, Y.; Whittall, R.M.; Li, L. Two-layer sample preparation: A method for MALDI-MS analysis of complex peptide and protein mixtures. *Anal. Chem.* **1999**, *71*, 1087–1091. [[CrossRef](#)] [[PubMed](#)]
24. Galleani, L.; Cohen, L.; Nelson, D. Local signal to noise ratio. In *Advanced Signal Processing Algorithms, Architectures, and Implementations XVI*; International Society for Optics and Photonics: Bellingham, WA, USA, 2006; Volume 6313, p. 63130Q. [[CrossRef](#)]
25. Mandrekar, J.N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. [[CrossRef](#)]
26. Dankers, F.J.W.M.; Traverso, A.; Wee, L.; van Kuijk, S.M.J. Prediction Modeling Methodology. In *Fundamentals of Clinical Data Science*; Springer International Publishing: Cham, Switzerland, 2019; pp. 101–120.
27. Baggerly, K.A.; Baggerly, K.A.; Coombes, K.R.; Morris, J.S. Bias, Randomization, and Ovarian Proteomic Data: A Reply to “Producers and Consumers”. *Cancer Inform.* **2005**, *1*, 9–14.
28. Li, K.; Pei, Y.; Wu, Y.; Guo, Y.; Cui, W. Performance of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) in diagnosis of ovarian cancer: A systematic review and meta-analysis. *J. Ovarian Res.* **2020**, *13*, 1–9. [[CrossRef](#)]
29. Alexandrov, T. MALDI imaging mass spectrometry: Statistical data analysis and current computational challenges. *BMC Bioinform.* **2012**, *13*, S11. [[CrossRef](#)]
30. Pan, S.; Zhang, H.; Rush, J.; Eng, J.; Zhang, N.; Patterson, D.; Comb, M.J.; Aebersold, R. High Throughput Proteome Screening for Biomarker Detection. *Mol. Cell. Proteomics* **2005**, *4*, 182–190. [[CrossRef](#)]
31. de Ridder, D.; de Ridder, J.; Reinders, M.J.T. Pattern recognition in bioinformatics. *Brief. Bioinform.* **2013**, *14*, 633–647. [[CrossRef](#)] [[PubMed](#)]
32. Swan, A.L.; Mobasher, A.; Allaway, D.; Liddell, S.; Bacardit, J. Application of Machine Learning to Proteomics Data: Classification and Biomarker Identification in Postgenomics Biology. *Omics A J. Integr. Biol.* **2013**, *17*, 595–610. [[CrossRef](#)] [[PubMed](#)]
33. Menon, U.; McGuire, A.J.; Raikou, M.; Ryan, A.; Davies, S.K.; Burnell, M.; Gentry-Maharaj, A.; Kalsi, J.K.; Singh, N.; Amso, N.N.; et al. The cost-effectiveness of screening for ovarian cancer: Results from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Br. J. Cancer* **2017**, *117*, 619–627. [[CrossRef](#)] [[PubMed](#)]
34. Jacobs, I.J.; Menon, U.; Ryan, A.; Gentry-Maharaj, A.; Burnell, M.; Kalsi, J.K.; Amso, N.N.; Apostolidou, S.; Benjamin, E.; Cruickshank, D.; et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): A randomised controlled trial. *Lancet* **2016**, *387*, 945–956. [[CrossRef](#)]
35. Iles, R.K.; Shahpari, M.E.; Cuckle, H.; Butler, S.A. Direct and rapid mass spectral fingerprinting of maternal urine for the detection of Down syndrome pregnancy. *Clin. Proteomics* **2015**, *12*, 9. [[CrossRef](#)]
36. Albrechtsen, J. Reproducibility in protein profiling by MALDI-TOF mass spectrometry. *Clin. Chem.* **2007**, *53*, 852–858. [[CrossRef](#)] [[PubMed](#)]
37. Antoniadis, A.; Bigot, J.; Lambert-Lacroix, S. Peaks detection and alignment for mass spectrometry data. *J. Société Française Stat.* **2010**, *151*, 17–37. [[CrossRef](#)]