International Journal of Bio-Science and Bio-Technology
Vol.5, No.6 (2013), pp.191-200
http://dx.doi.org/10.14257/ijbsbt.2013.5.6.20

Early and Advanced Stages of Ovarian Cancer

Hye-Jeong Song[1,3], Seung-Kyun Ko[2,3], Jong-Dae Kim[1,3],
Chan-Young Park[1,3] and Yu-Seop Kim[1,3]1

[1]*Dept. of Ubiquitous Computing, Hallym University, 1 Hallymdaehak-gil, Chuncheon,Gangwon-do, 200-702 Korea*
[2]*Dept. of Ubiquitous Game Engineering, Hallym University, 1 Hallymdaehak-gil, Chuncheon, Gangwon-do, 200-702, Korea*
[3]*Bio-IT Research Center, Hallym University, 1 Hallymdaehak-gil, Chuncheon, South Korea*

[2]*chokood@nate.com, { hjsong, kimjd, cypark, yskim01}@hallym.ac.kr*

Abstract

This study aimed to find optimal biomarker combinations (panels) by using logistic regression for the purpose of comparing the accuracy of distinguishing cancerous from benign tumors at the earlier (Stages I and II) and advanced stages (Stages III and IV) of ovarian cancer. Data samples were extracted from 120 patients with benign ovarian cysts and 65 patients with ovarian cancer. The concentrations of 21 urine biomarkers specific to ovarian cancer were obtained using Luminex and used in the study. Data samples were divided into early-stage and advantage-stage sample groups, and 2–3 biomarker combinations with the best area under the curve (AUC) values were selected for each stage using logistic regression to identify the optimal combinations for the accurate diagnosis of ovarian cancer. Additionally, AUC, sensitivity, specificity, diagnostic accuracy, and positive and negative predictive values of all of the selected biomarker combinations were compared. Among the combinations of 2 biomarkers, the best performing combination showed AUC values of 85.83% and 97.98% for the early and advanced stages, respectively; the same for the 3-biomarker combinations was 92.77% and 98.74%, respectively. These results confirm that in ovarian cancer diagnosis, biomarkers are more effective in early-stage detection than in advanced-stage detection.

Keywords: *Biomarker, Urine, Ovarian Cancer, Early Stages, Advanced Stages*

1. Introduction

Ovarian cancer commonly shows no perceivable symptoms in its early stages, and its early diagnosis is extremely difficult because non-cancerous and cancerous tumors are not easily distinguishable using radiation-free diagnostic tests such as ultrasonography. This often leads to costly and unnecessary surgical examinations. Epithelial ovarian cancer, which accounts for approximately 90% of ovarian cancer cases, is usually detected at Stage III or IV; therefore, once the cancer is diagnosed, the 5-year survival rate is <40%. For this reason, the

1 Corresponding author

ISSN: 2233-7849 IJBSBT

development of a diagnostic method capable of the early detection of ovarian cancer is of paramount importance [1-2].

Biomarker-based cancer diagnostic tests are relatively simple assays of the serum or urine, and the early diagnosis of cancer is possible through biomarkers at a lower cost than other screening and diagnostic tests [3].

The United States Food and Drug Administration (FDA) approved cancer-associated biomarkers for cancer diagnosis and issued related regulations and guidelines in the "In Vitro Diagnostic Multivariate Index Assay" (IVDMIA) document in 2007. According to the FDA definition, the IVDMIA is a technique that combines the values of multiple biomarkers to yield a single patient-specific interpretation function capable of deriving "classification," "score," and "index" for the purposes of disease diagnosis, cure, mitigation, treatment, and prevention [4].

Enhanced diagnostic accuracy using a method that combines several biomarkers and quantifies the analyses using statistical techniques is needed. The IVDMIA is significant since the availability of a single biomarker with a diagnostic power of quasi-100% specificity to any given type of cancer is currently lacking. The advantage of IVDMIA is the possibility of integrating information from complementary biomarkers while using a single index for a single marker, thus obtaining better outcomes than when they are used separately [5].

OVA1 [5], which was developed by Vermillion, Inc. as an ovarian cancer IVDMIA and approved by the FDA in 2009, is an ovarian cancer diagnostic blood test that uses 5 serum proteins as biomarkers. OVA1 is used to test patients with a pelvic mass requiring surgery and to diagnose whether a pelvic mass is cancerous or non-cancerous based on the test score on a scale of 0–10. Correlogic System, Inc. [6] developed a protein multivariate index-based ovarian cancer test using a microbead-based antibody assay, thus overcoming the diagnostic limitation of conventional biomarkers based on a single protein.

Amonkar [7] employed the random forest algorithm to discriminate ovarian cancer from a pelvic mass, in which the optimal combination of 11 biomarkers selected from 204 biomarkers showed a sensitivity of 85.7%. Brian [8] singled out the optimal subset from 55 biomarkers using the brand and bound algorithm to classify breast cancer and constructed a diagnostic model showing a sensitivity of 95% using a CART classification tree. Yurkovetsky [6] used the metropolis algorithm with Monte Carlo simulation to classify ovarian cancer and presented a classification model capable of finding 4 optimal biomarker combinations of 96 biomarkers. Borgia [9] evaluated 47 biomarkers using the area under the receiver operating characteristic curve (ROC AUC), selected 15 candidate markers, and ultimately identified 6 optimal marker combinations using the random forest algorithm, and the CART classification model built thereupon yielded 88% sensitivity.

Although most biomarker studies regarding ovarian cancer diagnostic tests use serum proteins, urine-based biomarkers are cancer diagnostic test methods that are completely non-surgical and easier to handle in clinical settings than serum-based biomarkers, and they can effectively provide timely differentiation of ovarian cancer from benign ovarian disease [10-12]. Petri *et al.*, [11] compared the ovarian cancer diagnosis-related serum and urine ROC AUC values of the same patient group and obtained the results of 84% and 83% for urine- and serum-based biomarkers, respectively, thus demonstrating little difference between them. The majority of women with clinical features corresponding to ovarian cancer are found to also have benign conditions. Therefore, the use of a method to differentiate ovarian cancer from benign conditions is meaningful [7].

Cancer staging is the process of determining the location and spread of cancer within the human body. Physicians establish treatment plans and make prognostic estimates for patients according to staging results [13].

Holschneider reported 5-year ovarian cancer Stages I, II, III, and IV survival rates of 93%, 70%, 37%, and 25%, respectively; in terms of early and advanced stages, the 5-year survival rates are 87% and 38%, respectively [14].

Shridhar also reported that while the 5-year survival rate of patients with Stage I cancer is >90%, that for patients with advanced stage cancer is <25% [15].

The present study aimed to compare the ability of classification models with high diagnostic accuracy to classify early and advanced stages using 21 types of urine biomarkers specific to ovarian cancer, and to identify the optimal marker combinations with the best classification performance for cancerous and non-cancerous ovarian tumors [9]. To achieve this, we created all possible combinations of 2–3 biomarkers from 21 biomarkers and assessed their ROC AUC values. The ovarian cancer diagnostic performance of the selected optimal biomarker combinations was then verified using logistic regression.

2. Data Set

The data samples were collected from 120 patients with benign ovarian cysts and 65 patients with ovarian cancer who underwent 21 types of urine biomarker tests. A total of 185 urine samples from Korean women were provided by ASAN Medical Center in Seoul. Table 1 lists the characteristics of the clinical samples.

Table 1. Information on the Clinical Samples Examined

Characteristics	No. of patients
No. of patients studied	185
Ovarian Cyst	120
Ovarian Cancer	65
Age (Mean ± S.D.)	45.37 ± 13.40
Range	21–82
FIGO stage	
Ⅰ	16 (24.61%)
Ⅱ	6 (9.23%)
Ⅲ	32 (49.23%)
Ⅳ	11 (16.92%)

The concentrations of urine protein biomarkers were measured using the multiplex immunoassay method with Luminex antibody microbeads; thereby, we used a multiplexed immunoassay kit consisting of cancer biomarkers specific to ovarian cancer [9]. Analyses were performed following the protocol of the manufacturer provided by Luminex Corp., and the samples were analyzed using the Bio-Plex Suspension Array System [16]. *The biomarker expression* levels *are shown* in terms of *median fluorescent intensities generated* from *analyzing microbeads in quantities of* 50–100 for an analyte of each sample. Analyte concentrations were quantified on the basis of the median fluorescent intensity using the standard curves generated by Bio-Rad (5-parameter curve fitting) [8].

3. Methods

Various variables should be calibrated at the time of urine collection prior to analysis. Calibration is performed to prevent concentration differences that dilute urinary protein using

creatinine as a reference point, which is relatively stable in urine [17].

To identify the best performing biomarker combinations for the differentiation of ovarian cancer from ovarian cysts, the classification performance of each biomarker combination subset should be measured until the best performing subset is found [18-20]. The cancer classification model should perform an adequate cancer-specific classification with respect to both sensitivity and specificity. A common method of verifying adequately high values of each consists of verifying the classification performance by measuring ROC AUC values [21, 22].

In this study, logistic regression was used to assess ROC AUC values to identify the biomarker combination with the highest value. To economize the biomarker selection time, the preliminary top 20 combinations were selected by means of 100 repetitions of 5-fold cross validations, and the second selection of top 3 combinations from the primarily selected ones was carried out by means of 1,000 repetitions of 5-fold cross validations, using mean AUC as a reference value. The cross-validation was performed 1000 times to reduce the deviation between the population and the samples. As illustrated in Figure 1, the AUC values were confirmed to converge when the cross-validation was repeated 1000 times.

Biomarker combination for the selection process was performed by establishing all possible combination of 2–3 biomarkers from 21 biomarkers, and the diagnostic performances of the selected biomarker combinations were evaluated against the logistic regression score threshold value of 0.5.

To assess bias due to the small sample size and use of the intermediate model, we used 5-fold bootstrap validation along with "out-of-bag" error estimation and 10% holdout bootstraps. Bootstrap estimates enable the evaluation of the model's potential value using only the study data. This method ensures the independence of the hold-out sample testing [7]. In the next stage, the early and advanced stages are compared using the classified biomarker combinations. Figure 2 shows the modeling process used to identify and classify the optimal biomarker combinations.

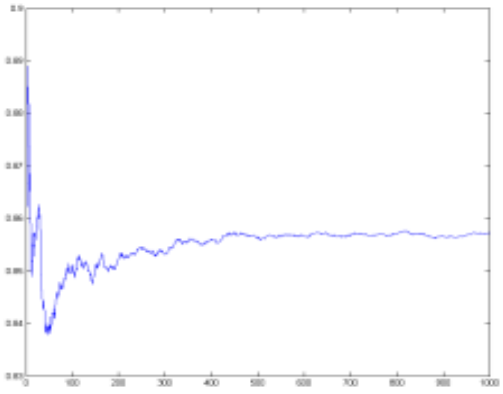

Figure 1. Area Under the Curve Convergence Graph

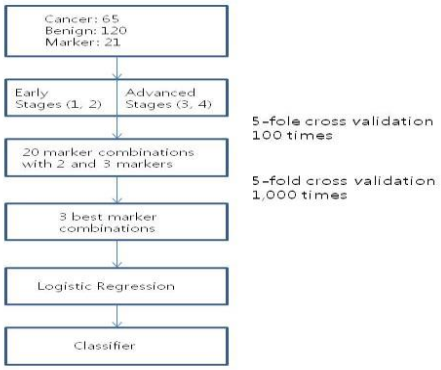

Figure 2. Ovarian Cancer Diagnostic Modeling Process

4. Results

Tables 2, 3, 4, and 5 outline the diagnostic performance test results of the optimal biomarker combinations obtained with the modeling. Tables 2 and 3 show the early- and advanced-stage diagnostic performance of the top three 2-biomarker combinations, while Tables 4 and 5 show the early- and advanced-stage diagnostic performance of the top three 3-biomarker combinations. A performance evaluation was conducted with the leave-one-out cross-validation method using a logistic regression classification algorithm with respect to AUC, sensitivity, specificity, classification accuracy, and positive and negative predictive values for classifying the early and advanced stages. In this study, the number of biomarkers used for combination was limited to 2 and 3 in an effort to limit test costs. The names of the biomarkers were not revealed to avoid infringing upon the patented biomarkers.

Table 2. Early-stage Diagnostic Performance (%) of top Three 2-biomarker Combinations

Marker	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
M5, M19	85.83	36.36	98.33	88.73	80	89.39
M5, M8	83.71	31.82	99.17	88.73	87.5	88.81
M5, M12	83.6	18.18	97.5	85.21	57.14	86.67

Table 3. Advanced-stage Diagnostic Performance (%) of top Three 2-biomarker Combinations

Marker	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
M5, M8	97.98	55.81	100	88.34	100	86.33
M4, M5	96.78	60.47	100	89.57	100	87.59
M5, M9	95.02	58.14	98.33	87.73	92.59	86.76

Table 4. Early-stage Diagnostic Performance (%) of Top Three 3-biomarker Combinations

Marker	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
M5, M8, M19	92.77	59.09	99.17	92.96	92.86	92.97
M5, M8, M12	91.67	45.45	99.17	90.85	90.91	90.84
M5, M8, M18	89.09	31.82	99.17	88.73	87.5	88.81

Table 5. Advanced-stage Diagnostic Performance (%) of Top Three 3-biomarker Combinations

Marker	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
M4, M5, M9	99.13	65.12	99.17	90.18	96.55	88.81
M4, M5, M19	98.74	67.44	100	91.41	100	89.55
M5, M8, M10	98.57	60.47	100	89.57	100	87.59

The best-performing 2-biomarker combinations for early-stage classification were M5 and M19 (AUC, 85.83%); those for advanced-stage classification were M5 and M8 (AUC, 97.98%). The best 3-biomarker combinations for early-stage classification were M5, M8, and M19 (AUC, 92.77%), while those for advanced-stage classification were M4, M5, and M9 (AUC, 99.13%). Figures 3, 4, 5, and 6 contain ROC diagrams that represent the data in Tables 2, 3, 4, and 5.

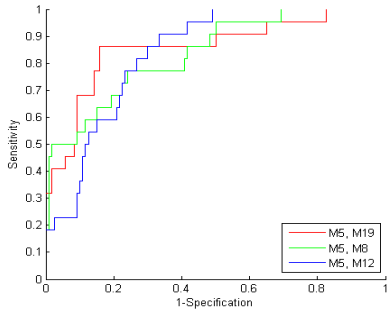

Figure 3. Early-stage ROC Curves of the Top Three 2-biomarker Combinations

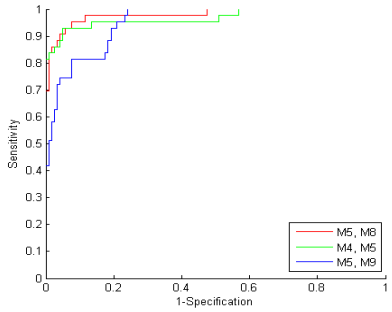

Figure 4. Advanced-stage ROC Curves of the Top Three 2-biomarker Combinations

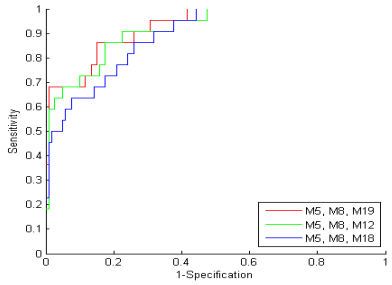

Figure 5. Early-stage ROC Curves of the Top three 3-biomarker Combinations

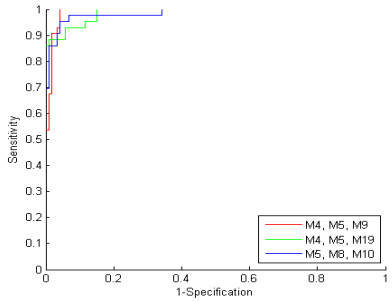

Figure 6. Advanced-stage ROC Curves of the Top Three 3-biomarker Combinations

5. Conclusion

In this study, we developed a novel diagnosis model utilizing multiple urinary protein-based biomarkers. A total of 21 types of urine biomarker specific to ovarian cancer were used to identify the best-performing biomarker combinations for the differentiation between ovarian cancer and ovarian cysts, and the diagnostic performance of each was verified. The optimal biomarker combinations were identified by assessing the ROC AUC values of all possible combinations of 2 or 3 biomarkers. The early- and advanced-stage classification performances of the biomarker combinations were then compared. The ovarian cancer diagnostic performance using 2-biomarker combinations improved to a greater degree for advanced-stage detection with a 97.98% AUC compared with an 85.83% AUC for early-stage detection. Similar values were confirmed for 3-biomarker combinations, with a 98.74% AUC for advanced-stage detection and a 92.77% AUC for early-stage detection.

Acknowledgement

The research was supported by the Research & Business Development Program through the Ministry of Knowledge Economy, Science and Technology (N0000425) and by Bio-IT Research Center, funded by Ahn-Gook Pharmaceutical Co., Ltd.

References

[1] American Cancer Society, http://www.cancer.org/Cancer/OvarianCancer, **(2012)**.
[2] N. Brian, M. Adele, V. Liudmila, P. Denise, W. Matthew, G. Elesier and L. Anna, "A Serum based analysis of ovarian epithelial tumorigenesis", ELSEVIER Gynecologic Oncology, vol. 112, **(2009)**, pp. 47-54.
[3] Q. Feng, M. Yu and N. B. Kiviat, "Molecular Biomarkers for Cancer Detection in Blood and Bodily Fluid", Critical Reviews in Clinical Laboratory Science, **(2006)**.
[4] US Department of Health and Human Services. Draft Guidance for Industry, Clinical Laboratories, and Staff: In Vitro Diagnostic Multivariate Index Assays, http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071455.pdf, **(2007)**.
[5] Z. Zhang, "An In Vitro Diagnostic Multivariate Index Assay(IVDMIA) for Ovarian Cancer: Harvesting the Power of Multiple Biomarkers", Rev Obstet Gynecol, vol. 5, no. 1, **(2012)**, pp. 35-41.
[6] Z. Yurkovetsky, S. Skates and A. Lomakin, "Development of a multimarker assay for early detection of ovarian cancer", J Clin Oncol., vol. 28, **(2010)**, pp. 2159-2166.
[7] S. D. Amonkar, G. P. Bertenshaw, T. H. Chen, K. J. Bergstrom, J. Zhao, P. Seshaiah, P. Yip and B. C. Mansfield, "Development and preliminary evaluation of a multivariate index assay for ovarian cancer", PLoS ONE, vol. 4, no. e4599, **(2009)**.
[8] B. M. Nolen, J. R. Marks, S. Ta'san, A. Rand, M. Luong, Y. Wang, K. Blackwell and A. E. Lokshin, "Serum biomarker profiles and response to neoadjuvant chemotherapy for locally advanced breast cancer", Breast Cancer Research, vol. 10, no. 3, **(2008)**.
[9] J. A. Borgia, S. Basu, L. P. Faber, A. W. Kim, J. S. Coon, K. A. Kaiser-alters, C. Fhied, S. Thomas, O. Rouhi, W. H. Warren, P. Bonomi and M. J. Liptay, "Establishment of a Multi-Analyte Serum Biomarker Panel to Identify Lymph Node Metastases in Non-small Cell Lung Cancer", Journal of Thoracic Oncology, vol. 4, no. 3, **(2009)**, pp. 338-347.
[10] B. M. Nolen and A. E. Lokshin, "Multianalyte assay systems in the differential diagnosis of ovarian cancer", Expert Opin. Med. Diagn., vol. 6, no. 2, **(2012)**, pp. 131-138.
[11] A. L. Petri, A. H. Simonsen, E. Høgdall, I. J. Christensen, S. K. Kjaer, C. Yip, S. Risum, A. T. Pedersen, D. Hartwell, E. T. Fung and C. Høgdall, "Comparison of proteomic biomarker panels in urine and serum for ovarian cancer diagnosis", Proteomics Clin Appl., vol. 4, no. 3, **(2010)**, pp. 304-314.
[12] Y. Kim, I. Koo, B. Hwa Jung, B. Chul Chung and D. Lee, "Multivariate classification of urine metabolome profiles for breast cancer diagnosis, BMC Bioinformatics", Molecular Biomarkers for Cancer Detection in Blood and Bodily Fluid, vol. 11, **(2010)**, pp. 1-9.
[13] American Cancer Society, http://www.cancer.org/treatment/understandingyourdiagnosis/staging.
[14] C. H. Holschneider and J. S. Berek, "Division of Gynecologic Oncology, Center for Health Sciences", UCLA School of Medicine, University of California, Los Angeles, California: Ovarian Cancer: Epidemiology, Biology, and Prognostic Factors, **(2000)**.
[15] V. Shridhar, J. Lee, A. Pandita, S. Iturruia, R. Avula, J. Staub, M. Morrissey, E. Calhoun, A. Sen, K. Kalli, G. Keeney, P. Roche, W. Cliby, K. Lu, R. Schmandt, G. B. Mills, R. C. Bast Jr., C. David James, F. J. Couch, L. C. Hartmann, J. Lillie and D. I. Smith, "Genetic Analysis of Ealry-versus Late-Stage Ovarian Tumors", **(2001)**.
[16] J. SunKyung, O. EunJi, Y. ChulWoo, A. WoongSik, K. YongGu, P. YeonJun and H. KyungJa, "ELISA for the selection of HLA isoantibody and Comparison Evaluation of Luminex Panel Reactive Antibody Test", Journal of Korean Society for Laboratory Medicine, vol. 29, **(2009)**, pp. 473-480.
[17] I. Hellstrom, P. J. Heagertty, E. M. Swisher, P. Liu, J. Jaffar, K. Agnew and K. Erik Hellstrom, "Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms", Cancer Lett., vol. 296, no. 1, **(2010)**, pp. 43-48.
[18] Y. Saeys, I. Inza and P. Larranaga, "A review of feature selection techniques in bioinformatics", Bioinformatics, vol. 19, **(2007)**, pp. 2507-2517.
[19] H.-J. Song, S.-K. Ko, J.-D. Kim, C.-Y. Park and Y.-S. Kim, "Looking for the Optimal Machine Learning Algorithm for the Ovarian Cancer Screening", International Journal of Bio-Science and Bio-Technology, vol. 5, no. 2, **(2013)**, pp. 41-48.
[20] Y.-S. Kim, M.-K. Jang, C.-Y. Park, H.-J. Song and J.-D. Kim, "Exploring Multiple Biomarker Combination by Logistic Regression for Early Screening of Ovarian Cancer", International Journal of Bio-Science and Bio-Technology, vol. 5, no. 2, **(2013)**, pp. 67-76.
[21] M. S. Pepe, T. Cai and G. L. Longton, "Combining predictors for classification using the area under the Receiver Operating Characteristic curve", Biometrics, vol. 62, **(2006)**, pp. 221-229.
[22] R. Pal, P. Garg, R. Chechi, S. Kumar and N. Kumar, "Cancer Growth Prediction via Artificial Neural Networks", International Journal of Bio-Science and Bio-Technology, vol. 2, no. 2, **(2010)**, pp. 1-10.

Authors

Hye-Jeong Song, she received the Ph.D. degree in Computer Engineering from Hallym University. She is a Professor in Department of Ubiquitous Computing, Hallym University. Her recent interests focus on biomedical system and bioinformatics

Seung-Kyun Ko, he is now the B.S. student in Ubiquitous Game Engineering of Hallym University. His recent interests focus on mobile game and bioinformatics

Jong-Dae Kim, he received the M.S. and the Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1984 and 1990, respectively. He worked for Samsung Electronics from 1988 to 2000 as an electrical engineer. He is a Professor in Department of Ubiquitous Computing, Hallym University. His recent interests focus on biomedical system and bioinformatics.

Chan-Young Park, he received the B.S. and the M.S. from Seoul National University and the Ph.D. degree from Korea Advanced Institute of Science and Technology in 1995. From 1991 to 1999, he worked at Samsung Electronics. He is currently a Professor in the Department of Ubiquitous Computing of Hallym University, Korea. His research interests are in Bio-IT convergence, Intelligent Transportation System and sensor networks.

Yu-Seop Kim, he received the Ph.D. degree in Computer Engineering from Seoul National University. He is currently a Professor in the Department of Ubiquitous Computing at Hallym University, South Korea. His research interests are in the areas of bioinformatics, computational intelligence and natural language processing.