Original Article
The clinical performance evaluation of novel protein chips for eleven biomarkers detection and the diagnostic model study

Yuan Luo1*, Xu Zhu2*, Pengjun Zhang3, Qian Shen3, Zi Wang1, Xinyu Wen4, Ling Wang5, Jing Gao5, Jin Dong5, Caie Yang1, Tangming Wu5, Zheng Zhu5, Yaping Tian1

1Core Laboratory of Translational Medicine, State Key Laboratory of Kidney Disease, Chinese PLA General Hospital, Beijing, P. R. China; 2Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Interventional Therapy Department, Peking University Cancer Hospital&Institute, Beijing, P. R. China; 3Department of Laboratory Diagnosis, Changhai Hospital, The Second Military Medical University, Shanghai, P. R. China; 4Department of Clinical Biochemistry, Chinese PLA General Hospital, Beijing, P. R. China; 5Jiangsu Sunlant Bioengineering Co., Ltd, Wuxi, P. R. China. *Co-first authors.

Received August 25, 2015; Accepted October 25, 2015; Epub November 15, 2015; Published November 30, 2015

Abstract: We aimed to develop and validate two novel protein chips, which are based on microarray chemiluminescence immunoassay and can simultaneously detected 11 biomarkers, and then to evaluate their clinical diagnostic value by comparing with the traditional methods. Protein chips were evaluated for limit of detection, specificity, common interferences, linearity, precision and accuracy. 11 biomarkers were simultaneously detected by traditional methods and protein chips in 3683 samples, which included 1723 cancer patients, 1798 benign diseases patients and 162 healthy controls. After assay validation, protein chips demonstrated high sensitivity, high specificity, good linearity, low imprecision and were free of common interferences. Compared with the traditional methods, protein chips have good correlation in the detection of all the 13 kinds of biomarkers (r≥0.935, P<0.001). For specific cancer detection, there were no statistically significant differences between the traditional method and novel protein chips, except that male protein chip showed significantly better diagnostic value on NSE detection (P=0.004) but significantly worse value on pro-GRP detection (P=0.012), female chip showed significantly better diagnostic value on pro-GRP detection (P=0.005). Furthermore, both male and female multivariate diagnostic models had significantly better diagnostic value than single detection of PGI, PG II, pro-GRP, NSE and CA125 (P<0.05). In addition, male models had significantly better diagnostic value than single CA199 and free-PSA (P<0.05), while female models observed significantly better diagnostic value than single CA724 and β-HCG (P<0.05). The new designed protein chips are simple, multiplex and reliable clinical assays and the multi-parameter diagnostic models based on them could significantly improve their clinical performance.

Keywords: Cancer biomarkers, protein array analysis, diagnostic tests, early detection of cancer, logistic models, ROC curve

Introduction
Nowadays, remarkable progresses have been made in understanding the molecular and cellular events that transform a normal functioning cell into a malignant one [1, 2]. Based on these progresses, cancer detection has entered into molecular diagnostic phase and the types of biomarkers are rapidly expanding, including proteins, peptides, hormones, certain DNA, microRNAs, and so on [3, 4]. Although it’s very significant to focus on the development of improved therapies for cancer patients, it’s still important to invent multiplex early detection methods, which could scent precancerous changes with high specificity and sensitivity [5]. Finding cancer at its earliest, most treatable stage gives patients the best chance to survive. Common methods recommended by American Cancer Society (ACS) for cancer early detection
Development, evaluation and model study of novel protein chips

Table 1. Characteristics of total studied subjects

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Malignant tumor (n=1723)</th>
<th>Benign diseases (n=1798)</th>
<th>Normal control (n=162)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>57.70</td>
<td>57.79</td>
<td>39.53</td>
</tr>
<tr>
<td>Range</td>
<td>9-91</td>
<td>9-97</td>
<td>13-54</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>837 (48.6)</td>
<td>822 (45.7)</td>
<td>97 (59.9)</td>
</tr>
<tr>
<td>Female</td>
<td>886 (51.4)</td>
<td>976 (54.3)</td>
<td>65 (40.1)</td>
</tr>
<tr>
<td>Tissue distribution, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>522 (30.3)</td>
<td>152 (8.5)</td>
<td>-</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>419 (24.3)</td>
<td>183 (10.2)</td>
<td>-</td>
</tr>
<tr>
<td>Liver and gall</td>
<td>192 (11.1)</td>
<td>226 (12.6)</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>590 (34.2)</td>
<td>1237 (68.8)</td>
<td>-</td>
</tr>
</tbody>
</table>

include many types, such as microbial detection, fecal occult blood test, digital rectal examination, double-contrast barium enema, computed tomography, tumor markers test, and so on [6]. Among these, the detection of tumor markers, which are directly produced or indirectly induced by the tumor cell and could reflect its growth and/or activity, is one of the most convenient means to monitor the insignificant or potential cancer transformation [7].

However, tumor markers are not specific to cancer as their name looks like and the majority of them could be synthesized and released by normal cells, too. Consequently, the specificity of tumor markers is not their presence at tumor sites but the concentrations detected there, which is why normal values are needed [8]. Moreover, present clinical proven methods, such as radioimmunoassay or chemiluminescence, could only detect one biomarker at a time. In other words, they don’t observe high-throughput abilities, which result in low working efficiency and suppress their use in mass screening. On the contrary, protein array technology possesses the greatest potential for screening thousands of recombinant- or purified proteins, as well as synthetic peptides and even unknown proteins in a high throughput [9, 10]. Furthermore, since various benign pathologies could also produce tumor markers, which give rise to false positives, single tumor marker has relatively low specificity and sensitivity and might sometimes shows equivocal results. Therefore, how to improve the detection efficiency and increase the specificity and sensitivity is the most important aspect of diagnostic methods [11-12]. To break through this bottleneck, using multiple biomarkers or indicators to establish diagnostic models is a common effective strategy, which could integrate the advantages of each index and simultaneously improve both sensitivity and specificity [13-15].

In our study, we have developed two protein chips, which are based on micro-array chemiluminescence immunoassay and could simultaneously detect 11 kinds of biomarkers of man or woman, respectively. Specifically, Pepsinogen I (PGI), Pepsinogen II (PGII), pro-Gastrin-releasing peptide (pro-GRP), alpha-fetoprotein (AFP), Neuron-specific enolase (NSE), Carcinoembryonic antigen (CEA), Carbohydrate antigen 125 (CA125), Carbohydrate antigen 199 (CA199) and Carbohydrate antigen 724 (CA724) are common in both male and female chips. Meanwhile, total-Prostate specific antigen (t-PSA) and free-Prostate specific antigen (f-PSA) are unique for male chip, beta-Human chorionic gonadotropin (β-HCG) and Carbohydrate antigen 153 (CA153) are unique for female chip. Then, protein chips were evaluated for limit of detection, specificity, common interferences, linearity, precision and accuracy. After that, diagnostic models were established to improve the detection specificity and sensitivity at the same time. We aimed to evaluate the clinical performance of the protein chips and the feasibility of using multiple biomarkers to establish diagnostic models for cancer detection.

Patients and methods

Patients and healthy controls

Our study was approved by ethics committee of the Chinese PLA General Hospital, Shanghai Changzheng Hospital, Nanjing general hospital of Nanjing military command and the second affiliated hospital of Zhejiang university school of medicine and has been performed in accordance with the ethical standards laid down in an appropriate version of the 2000 Declaration of Helsinki as well as the Declaration of Istanbul.
Figure 1. Detection correlation between conventional method and protein chips. New- represents results detected by protein chips, while old- represents results detected by conventional method.
Development, evaluation and model study of novel protein chips

Figure 2. Receiver operating characteristic curve analysis of biomarkers detected by conventional methods, male chip or multivariate diagnostic models for evaluating the diagnostic value of specific cancer prediction. *P is the p value of z statistic compared between conventional methods and male chip, while #p is the p value of that between conventional method and multivariate diagnostic model.
**Figure 3.** Receiver operating characteristic curve analysis of biomarkers detected by old methods, female chip or multivariate diagnostic models for evaluating the diagnostic value of specific cancer prediction. *P* is the *p* value of *z* statistic compared between conventional methods and female chip, while #p is the *p* value of that between conventional method and multivariate diagnostic model.
2009. With written consent, the serum samples were collected and detected in four hospitals mentioned above. All the four hospitals used the same batch protein chip detection reagent and machine to ensure that the results were reliable and comparable. 1723 cancer patients, 1798 benign disease patients and 162 healthy controls were enrolled in the study. Both of the cancer and benign patients’ samples were collected before any treatment, such as surgery, chemotherapy or radiation therapy, and further confirmed by histopathological analysis. Healthy controls were enrolled based on their negative results on the blood biomarker test, x-ray, ultrasound, computed tomography examination and fecal occult blood testing. The clinical characteristics of the participants were shown in Table 1.

Serum collection

A total of 5 ml of peripheral blood samples was collected in tubes containing a separating gel and clot activator. After blood collection, the tubes should be upside down ten times quickly and then rest for about 1 hour. After centrifuging at 2500 rpm for 5 min, the samples could be directly detected by corresponding techniques. If it’s not convenient to do the test at once, the supernatant could be transferred into new tubes, and the serum should be stored in aliquots at -80°C until detection.

Measurement of serum biomarker levels using novel protein chips

The microarray-based sandwich immunoassay kits (Sunlant Bioengineering, Wuxi, China), which are cooperated with protein chips, were used to detect the serum concentrations of cancer related biomarkers. Briefly, serum sample and microarray that was covalently coupled to specific capture antibodies were incubated together. After washing, a cocktail of horseradish peroxidase-labeled antibodies were added to the array, which was shaken at room temperature. Then, the array was transferred to a cuvette with substrate solution and detected by an automatic biochip reader. The results were recorded by standard curves established by five different concentrations.

Measurement of serum biomarker levels using traditional detection methods

PGI, PGII and pro-GRP were detected by an Abbott i2000 instrument (Abbott Laboratories, USA). AFP, NSE, CEA, CA125, CA199, CA724, PSA, f-PSA, β-HCG and CA153 were detected by Roche Cobas e601 instrument (Roche Diagnostics, Germany). The main principle of these methods is competitive protein binding (CPB), however they have different detection methods: electrochemiluminescence detection in Roche Cobas e601 and chemiluminescence detection in Abbott i2000. Detection was conducted according to the manufacturer’s reagent kit manual for specific biomarkers.

Evaluation of detection sensitivity and specificity of protein chips

For sensitivity evaluation, the experiment to determine the limit of detection (LOD) was performed on 3 separate occasions, as follows: the zero concentration calibrator was evaluated for 15 replicates, respectively. The average and 2SDs (Standard deviation) of the determinations was used to calculate LOD. For specificity evaluation, 50 samples of patients with high expression of various biomarkers were detected by protein chips to validate their detection specificity. In addition, 30 samples from hemoysis, hyperlipidemia or jaundice were used to validate the detection stability.

Evaluation of detection linearity of protein chips

For each biomarker, linearity was evaluated using Abbott or Roche calibrators in triplicate measurements. Linearity was also evaluated using a patient sample with high concentration of specific biomarker. This sample was diluted with blank whole blood at different proportions to achieve 7 dilutions. These diluted specimens were each examined in triplicate.

Evaluation of precision performance of protein chips

According to NCCLS EP5-A program [16], both intra- and inter-batch examination used the same machine and standard curve but three reagent box with different lot number. For the former, repeat detecting two internal quality controls for 20 times in one day and then calculating the intra-batch precision. For the latter, repeat detecting one internal quality control for 2 times in one day and continuously monitoring for 20 days and then calculating the inter-batch precision.
Statistical analyses

Statistical analyses were carried out using SPSS 22.0 (SPSS Inc., USA) and MedCalc 13.0 (MedCalc Software, Belgium). Pearson’s correlation coefficient was used for correlation analysis of the associated quantitative data. Kappa test was used to detect the diagnostic consistency between methods. Receiver operating characteristic (ROC) curves were established to evaluate the diagnostic value, and the Areas under the ROC curves (AUC) were compared by z-scores test. The cutoff value, which determined the sensitivity and specificity, was chosen by the Youden index. Binary logistic regression models were built to evaluate the multivariate diagnostic value of multiple biomarkers detected by protein chips. A two-tailed p-value of less than 0.05 was considered statistically significant.

Results

Evaluation of detection sensitivity and specificity of protein chips

The LOD for protein chips were defined here as the determination at 2SDs above the zero concentration calibrators. Based on 3 separate analyses, the average LOD for PGI, PGII, pro-GRP, AFP, NSE, CEA, CA125, CA199, CA724, t-PSA, f-PSA, β-HCG and CA153 was 4 ng/ml, 1.5 ng/ml, 0.02 ng/ml, 2 ng/ml, 2.5 ng/ml, 0.5 ng/ml, 2.3 U/ml, 5 U/ml, 1 U/ml, 0.1 ng/ml, 0.1 ng/ml, 0.4 ng/ml and 5 U/ml. Meanwhile, protein chips could specifically detect every biomarker on the chips without cross-reactivity. Additionally, interference potential was evaluated by determining biomarker concentration for samples from hemolysis, hyperlipidemia or jaundice. The results showed that severe hemolysis samples are prone to lead false positive of NSE index, while samples from hyperlipidemia or jaundice have no impacts on the chip results (data not shown). It suggested that protein chips had comparable detection sensitivity and specificity with traditional methods and could meet the basic requirements of clinical use.

Evaluation of detection linearity of protein chips

After linearity evaluation with Abbot or Roche calibrators and serial dilutions of patient samples, protein chips were found to be linear to detect PGI, PGII, pro-GRP, AFP, NSE, CEA, CA125, CA199, CA724, t-PSA, f-PSA, β-HCG and CA153 in the range of 6-300 ng/ml, 2-200 ng/ml, 0.03-5 ng/ml, 3-1000 ng/ml, 2.6-300 ng/ml, 1-300 ng/ml, 3.5-3000 U/ml, 7-2000 U/ml, 1.35-400 U/ml, 0.2-100 ng/ml, 0.6-200 ng/ml and 7-300 U/ml. For every biomarker, the correlation coefficient of linearity was higher than 0.99 (r≥0.99). It suggested that protein chip had good detection linearity and wide linear range.

Evaluation of precision performance of novel protein chips

In order to evaluate the precision performance, protein chips for male and female were respectively used to do the intra- and inter-batch examination. As shown in Supplementary Tables 1, 2, 3 and 4, coefficient of variation (%CV) of male chip in intra-batch examination ranged from 2.4% to 10.0%, %CV of male chip in inter-batch examination ranged from 5.0% to 9.1%, %CV of female chip in intra-batch examination ranged from 1.89% to 9.95%, %CV of female chip in inter-batch examination ranged from 4.66% to 9.56%. It suggested that both male and female protein chips have good detection precision in intra- and inter-batch. In other words, protein chips are reliable detection methods and had good repeatability and stability.

Correlation analysis between the traditional methods and novel protein chips

All kinds of biomarkers were detected by both the traditional methods and novel protein chips, and then the same biomarkers from both chips were analyzed together. As showed in Figure 1, the correlation coefficient for PGI, PGII, pro-GRP, AFP, NSE, CEA, CA125, CA199, CA724, t-PSA, f-PSA, β-HCG and CA153 was 0.935, 0.947, 0.988, 0.985, 0.969, 0.980, 0.980, 0.980, 0.978, 0.979, 0.982, 0.981 and 0.966. And all the p values for them were 0.000. It suggested that protein chips have good correlation and detection consistency with the traditional method in the detection of all the 13 kinds of biomarkers.

Diagnostic value of the 11 biomarkers detected by traditional methods and novel protein chips for specific cancers detection

After the assay validation and correlation analysis, we examined the diagnostic ability of the
Development, evaluation and model study of novel protein chips

biomarkers depending on their clinical utility. Generally, PGI, PGII, CEA, CA199 and CA724 were used for gastrointestinal tumors detection, while pro-GRP and NSE for lung cancer, AFP for liver cancer, t-PSA and f-PSA for prostate cancer, β-HCG for gynecologic tumor, CA153 for female breast cancer, and CA125 for all cancers. The diagnostic value of the two methods on specific cancer screening was evaluated by ROC curves and their AUCs were compared by z-score test. Benefit from the multiplex detection ability of protein chips, we further established binary logistic regression models to evaluate their multivariate diagnostic value in specific cancer screening. The diagnostic models and their specificity (SPC) and sensitivity (SEN) were listed in Supplementary Table 5.

As shown in Figures 2, 3, there were no statistically significant differences between the traditional method and novel protein chips, except that male protein chip showed significantly better diagnostic value on NSE detection (P=0.004) but significantly worse value on pro-GRP detection (P=0.012), female chip showed significantly better diagnostic value on pro-GRP detection (P=0.005). Furthermore, compared with traditional methods, multivariate diagnostic models

Figure 4. Receiver operating characteristic curve analysis of multivariate diagnostic models for evaluating the diagnostic value of disease or cancer differentiation. The sensitivity and specificity were determined by the maximum Youden index. AUC: Areas under the curve.
Development, evaluation and model study of novel protein chips

established on chip results had significantly better diagnostic value as a whole. Specifically, both male and female models had significantly better diagnostic value than single detection of PGI, PG II, pro-GRP, NSE and CA125 (P<0.05). In addition, male models had significantly better diagnostic value than single CA199 and free-PSA (P<0.05), while female models observed significantly better diagnostic value than single CA724 and β-HCG (P<0.05). It suggested that protein chips are comparable with traditional methods in single biomarker detection manner. Once establishing diagnostic models based on multiplex chip results, it observed significantly higher diagnostic value than single detection of 7 biomarkers in 11. It showed that protein chips could not only increase the detection efficiency, but also improve the diagnostic value a lot.

Multivariate diagnostic value of protein chips for total disease or cancer detection

Similarly, based on the multiplex detection ability of protein chips, we have established binary logistic regression models to evaluate their multivariate diagnostic value in total disease or cancer detection. Specifically, in total disease identification, both the cancer and benign disease patients were considered as “disease positive” group, while the normal control as “disease negative” group. Meanwhile, in the case of total cancer identification, the cancer patients were set as “cancer positive” group, while the benign disease patients as “cancer negative” group. The diagnostic models and their SPC and SEN were listed in Supplementary Table 6. As shown in Figure 4, the AUC of multivariate logistic regression for the male and female disease detection was 0.981 (95% CI: 0.975-0.987) and 0.836 (95% CI: 0.798-0.874), respectively. While, that for cancer detection was 0.691 (95% CI: 0.666-0.717) and 0.753 (95% CI: 0.731-0.775), respectively. The results preliminarily indicated that protein chips might have a certain diagnostic value in the detection of total diseases or tumors.

Discussion

Cancer, which is one of the most common causes of disease-associated death, is a major health issue in the world [17]. Fortunately, a substantial proportion of cancers, which are caused by cigarette smoking, heavy use of alcohol or other bad habits, could be prevented if people could reduce such harmful exposure. Moreover, some cancers could even be cured if people could scent precancerous changes and get diagnosed at an early stage, such as cancers of the breast, cervix, colon, and rectum [18-20]. Therefore, early detection and timely diagnosis is the key point of cancer prevention and control, which also predicts good clinical outcomes. Unexpectedly, although developed countries could provide the best services for detection, diagnosis, and treatment, they have the highest incidence of cancer. The highest prevalence proportions of cancer also occur in these populations. This could be explained by the following three reasons. Firstly, developed countries observe high level of medical technology, which could detect more precancerous patients at early stages, while these patients often miss their survival chances in developing countries. Secondly, most developed countries are aging society, which has a higher incidence of cancer. At last, higher incidence of cancer often means that other diseases, such as infectious ones, are less. So, this unexpected incidence level doesn’t mean that there’s no need to achieve more abilities of cancer detection, diagnosis and treatment.

As reported by Chinese National Cancer Institute, lung cancer, gastric cancer, liver cancer, esophageal cancer, colorectal cancer, female breast cancer and cervical cancer were the most common cancers in China [21]. In our study, various biomarkers most related to the aforementioned cancers were detected by clinical proven methods and new designed protein chips, named male chip and female chip, which are based on microarray chemiluminescent immunoassay and could simultaneously detect 11 kinds of biomarkers. Mukoubayashi C and his colleagues have found that serum PGI/II testing is useful in detecting early-stage gastric cancers arising from atrophic gastric mucosa [22]. As reviewed by Ischia J, pro-GRP is the most sensitive serum biomarker in patients with small cell lung cancer and is a significant predictor of poor survival in patients with advanced prostate cancer [22]. Interestingly, β-HCG is not only an important hormone, but also could be used as tumor marker for some gynecological cancers [24]. Although with various levels of SPC and SEN, the rest of the biomarkers targeted by protein chips are common clinical proven tumor markers [25-27].
After comprehensive assay validation, protein chips demonstrated high sensitivity, high specificity, good linearity, low imprecision and were free of common interferences. Compared with the traditional methods, protein chips have good correlation in the detection of all the 13 kinds of biomarkers (r≥0.935, P<0.001). As detailed in the results, we further compared the diagnostic ability among traditional methods, protein chips and diagnostic models based on the multiplex chip results. The ROC curve results suggest that both traditional methods and protein chips are roughly the same in detecting single biomarker for cancer detection, except NSE and pro-GRP. Specifically, female chip observed better diagnostic value in predicting lung cancer through pro-GRP detection, while male chip had worse value of that. In addition, male chip had better diagnostic value in NSE detection for lung cancer discrimination. Once establishing diagnostic models based on multiplex chip results, it could significantly enhance their diagnostic ability. Thereafter, trying to improve the detection specificity and sensitivity of male or female chip, we have established 4 diagnostic formulas based on multivariate logistic regression analyses for total disease or cancer detection, respectively. Considering the SPC and SEN, it showed that protein chips might be useful in total disease detection but play limited roles in cancer discrimination. However, due to the small ratio of normal control in both male (97 in 1756) and female (65 in 1927) group, the high SPC of total disease differentiation might not be as real as it showed.

In conclusion, we have comprehensively assessed the new designed protein chips, which could simultaneously detect 11 kinds of biomarkers closely related to the most common cancers in China. After assay validation and diagnostic value comparison, we found that the new designed protein chips have generally equal diagnostic ability to the traditional methods and they observe high-throughput detection capabilities. Furthermore, combining with multivariate logistic regression, we found that diagnostic models established from the chip results might be useful in cancer detection. In a word, these chips might play important roles in cancer early detection, especially for mass screening, which has urgent need for high-throughput testing methods. Although this study was conducted in four different hospitals, the sample size was still small, especially that of the healthy control group. Besides, the sample size of specific diseases or cancer was also small. Hence, multicenter studies with larger sample sizes will be needed to further validate the value of the chips in diagnosing cancer.

**Acknowledgements**

This study was supported by the National High Technology Research and Development Program of China (863 Program, No. 2011AA02-A111), the National Natural Science Foundation of China (No. 81371906) and the Capital Health Development Special Scientific Research Projects (2014-2-2154). And we wish to thank all of the doctors and nurses in the above four mentioned hospitals for their kindly support and assistance.

**Disclosure of conflict of interest**

None.

**Address correspondence to:** Yaping Tian, Core Laboratory of Translational Medicine, State Key Laboratory of Kidney Disease, Chinese PLA General Hospital, Beijing, P. R. China. E-mail: tianyp61@gmail.com

**References**


Development, evaluation and model study of novel protein chips


## Development, evaluation and model study of novel protein chips

### Supplementary Table 1. Intra-batch evaluation of precision performance of male chip

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Sample</th>
<th>Index</th>
<th>PGI (ng/ml)</th>
<th>PGII (ng/ml)</th>
<th>CA199 (U/ml)</th>
<th>AFP (ng/ml)</th>
<th>PSA (ng/ml)</th>
<th>F-PSA (ng/ml)</th>
<th>NSE (ng/ml)</th>
<th>CA125 (U/ml)</th>
<th>CEA (ng/ml)</th>
<th>CA724 (U/ml)</th>
<th>Pro-GRP (ng/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lot 1 Internal quality control 1</td>
<td>Mean</td>
<td>66.3</td>
<td>6.0</td>
<td>113.3</td>
<td>328.6</td>
<td>5.5</td>
<td>3.1</td>
<td>13.8</td>
<td>141.2</td>
<td>23.6</td>
<td>19.9</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.6</td>
<td>0.2</td>
<td>9.8</td>
<td>24.2</td>
<td>0.3</td>
<td>0.3</td>
<td>1.3</td>
<td>5.4</td>
<td>1.9</td>
<td>1.8</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>3.9</td>
<td>4.1</td>
<td>8.6</td>
<td>7.4</td>
<td>5.0</td>
<td>9.4</td>
<td>9.1</td>
<td>3.8</td>
<td>8.2</td>
<td>8.9</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>329.3</td>
<td>22.9</td>
<td>526.9</td>
<td>931.9</td>
<td>25.1</td>
<td>16.4</td>
<td>62.6</td>
<td>746.2</td>
<td>91.3</td>
<td>99.3</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>7.8</td>
<td>1.6</td>
<td>50.5</td>
<td>30.8</td>
<td>2.2</td>
<td>1.5</td>
<td>6.1</td>
<td>34.8</td>
<td>7.9</td>
<td>7.5</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>2.4</td>
<td>6.9</td>
<td>9.6</td>
<td>3.3</td>
<td>8.8</td>
<td>8.9</td>
<td>9.7</td>
<td>4.7</td>
<td>8.7</td>
<td>7.6</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>Lot 2 Internal quality control 1</td>
<td>Mean</td>
<td>70.5</td>
<td>7.5</td>
<td>130.3</td>
<td>361.2</td>
<td>5.3</td>
<td>2.9</td>
<td>13.8</td>
<td>136.3</td>
<td>22.6</td>
<td>20.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.2</td>
<td>0.3</td>
<td>10.7</td>
<td>19.9</td>
<td>0.2</td>
<td>0.2</td>
<td>1.3</td>
<td>5.7</td>
<td>2.1</td>
<td>1.3</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>3.1</td>
<td>3.8</td>
<td>8.2</td>
<td>5.5</td>
<td>3.9</td>
<td>8.1</td>
<td>9.1</td>
<td>4.2</td>
<td>9.3</td>
<td>6.4</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>371.1</td>
<td>2.2</td>
<td>553.4</td>
<td>1021.9</td>
<td>26.7</td>
<td>16.4</td>
<td>64.9</td>
<td>684.5</td>
<td>97.7</td>
<td>93.8</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>27.0</td>
<td>0.1</td>
<td>49.5</td>
<td>56.7</td>
<td>1.9</td>
<td>1.5</td>
<td>6.0</td>
<td>29.3</td>
<td>6.7</td>
<td>7.0</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>7.3</td>
<td>6.5</td>
<td>8.9</td>
<td>5.6</td>
<td>7.1</td>
<td>9.2</td>
<td>9.3</td>
<td>4.3</td>
<td>6.9</td>
<td>7.5</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Lot 3 Internal quality control 1</td>
<td>Mean</td>
<td>72.3</td>
<td>7.6</td>
<td>113.3</td>
<td>318.1</td>
<td>5.2</td>
<td>2.7</td>
<td>13.9</td>
<td>137.4</td>
<td>24.2</td>
<td>19.0</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.9</td>
<td>0.3</td>
<td>9.8</td>
<td>21.6</td>
<td>0.2</td>
<td>0.3</td>
<td>1.4</td>
<td>5.1</td>
<td>2.3</td>
<td>1.7</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>4.0</td>
<td>3.6</td>
<td>8.6</td>
<td>6.8</td>
<td>3.3</td>
<td>9.6</td>
<td>10.0</td>
<td>3.7</td>
<td>9.5</td>
<td>8.7</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>395.5</td>
<td>2.1</td>
<td>537.4</td>
<td>960.1</td>
<td>27.7</td>
<td>13.9</td>
<td>67.3</td>
<td>681.5</td>
<td>99.7</td>
<td>91.6</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>21.6</td>
<td>0.2</td>
<td>38.4</td>
<td>80.8</td>
<td>1.3</td>
<td>1.3</td>
<td>6.3</td>
<td>21.1</td>
<td>9.8</td>
<td>6.7</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CV, %</td>
<td>5.5</td>
<td>8.7</td>
<td>7.2</td>
<td>8.4</td>
<td>4.8</td>
<td>9.0</td>
<td>9.4</td>
<td>3.1</td>
<td>9.9</td>
<td>7.3</td>
<td>6.7</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Table 2. Inter-batch evaluation of precision performance of male chip

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Reagent</th>
<th>Precision among days</th>
<th>Inter-batch precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>CA199</td>
<td>Lot 1</td>
<td>111.3</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>111.7</td>
<td>10.1</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>110.3</td>
<td>9.3</td>
</tr>
<tr>
<td>AFP</td>
<td>Lot 1</td>
<td>161.5</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>162.7</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>164.2</td>
<td>9.7</td>
</tr>
<tr>
<td>PGI</td>
<td>Lot 1</td>
<td>66.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>67.4</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>68.4</td>
<td>4.7</td>
</tr>
<tr>
<td>PGIi</td>
<td>Lot 1</td>
<td>4.8</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>4.9</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>4.9</td>
<td>0.4</td>
</tr>
<tr>
<td>CEA</td>
<td>Lot 1</td>
<td>23.0</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>22.8</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>22.6</td>
<td>1.4</td>
</tr>
<tr>
<td>CA125</td>
<td>Lot 1</td>
<td>146.7</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>144.9</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>145.4</td>
<td>13.0</td>
</tr>
<tr>
<td>NSE</td>
<td>Lot 1</td>
<td>13.5</td>
<td>0.941</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>13.2</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>13.2</td>
<td>0.996</td>
</tr>
<tr>
<td>Pro-GRP</td>
<td>Lot 1</td>
<td>0.4</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>0.4</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>0.4</td>
<td>0.03</td>
</tr>
<tr>
<td>T-PSA</td>
<td>Lot 1</td>
<td>5.7</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>5.7</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>5.8</td>
<td>0.4</td>
</tr>
<tr>
<td>F-PSA</td>
<td>Lot 1</td>
<td>3.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>2.9</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>2.9</td>
<td>0.2</td>
</tr>
<tr>
<td>CA724</td>
<td>Lot 1</td>
<td>19.3</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>19.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>19.5</td>
<td>1.4</td>
</tr>
</tbody>
</table>
### Supplementary Table 3. Inter-batch evaluation of precision performance of female chip

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Sample</th>
<th>Index</th>
<th>PG1 ng/ml</th>
<th>PG2 ng/ml</th>
<th>CA199 U/ml</th>
<th>AFP ng/ml</th>
<th>CA153 ng/ml</th>
<th>β-HCG ng/ml</th>
<th>NSE ng/ml</th>
<th>CA125 U/ml</th>
<th>CEA ng/ml</th>
<th>CA724 U/ml</th>
<th>Pro-GRP ng/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lot 1</td>
<td>Internal quality control 1</td>
<td>Mean</td>
<td>70.7</td>
<td>6.91</td>
<td>120.26</td>
<td>313.86</td>
<td>9.69</td>
<td>3.88</td>
<td>1.29</td>
<td>142.8</td>
<td>23.57</td>
<td>20.39</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>2.26</td>
<td>0.23</td>
<td>9.41</td>
<td>18.50</td>
<td>0.79</td>
<td>0.35</td>
<td>0.12</td>
<td>7.26</td>
<td>2.10</td>
<td>1.8</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>3.19</td>
<td>3.34</td>
<td>7.82</td>
<td>5.89</td>
<td>8.11</td>
<td>8.99</td>
<td>9.00</td>
<td>5.08</td>
<td>8.92</td>
<td>8.83</td>
<td>5.69</td>
</tr>
<tr>
<td></td>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>359.3</td>
<td>21.18</td>
<td>280.36</td>
<td>414.55</td>
<td>41.26</td>
<td>28.75</td>
<td>13.88</td>
<td>741</td>
<td>86.03</td>
<td>102.3</td>
<td>1.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>17.76</td>
<td>1.195</td>
<td>15.87</td>
<td>15.38</td>
<td>1.58</td>
<td>2.7</td>
<td>1.07</td>
<td>29.85</td>
<td>3.81</td>
<td>6.85</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>4.94</td>
<td>5.64</td>
<td>5.66</td>
<td>3.71</td>
<td>3.83</td>
<td>9.38</td>
<td>7.74</td>
<td>4.03</td>
<td>4.43</td>
<td>6.70</td>
<td>4.97</td>
</tr>
<tr>
<td>Lot 2</td>
<td>Internal quality control 1</td>
<td>Mean</td>
<td>66.88</td>
<td>7.35</td>
<td>108.75</td>
<td>314.29</td>
<td>8.72</td>
<td>3.66</td>
<td>13.01</td>
<td>132.7</td>
<td>22.95</td>
<td>18.84</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>2.155</td>
<td>0.27</td>
<td>8.07</td>
<td>18.14</td>
<td>0.77</td>
<td>0.31</td>
<td>1.16</td>
<td>6.96</td>
<td>2.28</td>
<td>1.41</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>3.22</td>
<td>3.64</td>
<td>7.42</td>
<td>5.77</td>
<td>8.77</td>
<td>9.10</td>
<td>8.92</td>
<td>5.25</td>
<td>9.95</td>
<td>7.48</td>
<td>3.43</td>
</tr>
<tr>
<td></td>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>236</td>
<td>16.74</td>
<td>235.45</td>
<td>415.21</td>
<td>41.37</td>
<td>26.13</td>
<td>116.4</td>
<td>714.1</td>
<td>79.41</td>
<td>95.75</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>4.46</td>
<td>0.54</td>
<td>15.7</td>
<td>21.53</td>
<td>1.96</td>
<td>2.41</td>
<td>6.35</td>
<td>38.1</td>
<td>4.31</td>
<td>7.29</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>1.89</td>
<td>3.24</td>
<td>6.67</td>
<td>5.19</td>
<td>4.74</td>
<td>9.22</td>
<td>5.45</td>
<td>5.34</td>
<td>5.43</td>
<td>7.61</td>
<td>4.09</td>
</tr>
<tr>
<td>Lot 3</td>
<td>Internal quality control 1</td>
<td>Mean</td>
<td>68.54</td>
<td>6.97</td>
<td>118.85</td>
<td>328.6</td>
<td>9.32</td>
<td>3.56</td>
<td>13.58</td>
<td>131.3</td>
<td>22.98</td>
<td>19.35</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>2.65</td>
<td>0.21</td>
<td>10.46</td>
<td>24.22</td>
<td>0.79</td>
<td>0.32</td>
<td>1.24</td>
<td>4.98</td>
<td>2.18</td>
<td>1.79</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>3.87</td>
<td>2.98</td>
<td>8.80</td>
<td>7.37</td>
<td>8.50</td>
<td>8.83</td>
<td>9.14</td>
<td>3.80</td>
<td>9.47</td>
<td>9.25</td>
<td>4.21</td>
</tr>
<tr>
<td></td>
<td>Internal quality control 2</td>
<td>Mean</td>
<td>231.7</td>
<td>15.08</td>
<td>243.73</td>
<td>424.51</td>
<td>42.64</td>
<td>24.97</td>
<td>123.8</td>
<td>718.9</td>
<td>90.59</td>
<td>97.07</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>12</td>
<td>0.55</td>
<td>19.24</td>
<td>22.54</td>
<td>1.62</td>
<td>1.81</td>
<td>7.12</td>
<td>30.99</td>
<td>2.0</td>
<td>7.29</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CV, %</td>
<td>5.18</td>
<td>3.67</td>
<td>7.89</td>
<td>5.31</td>
<td>3.81</td>
<td>7.26</td>
<td>5.75</td>
<td>4.31</td>
<td>2.21</td>
<td>7.51</td>
<td>5.51</td>
</tr>
</tbody>
</table>
## Supplementary Table 4. Inter-batch evaluation of precision performance of female chip

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Reagent</th>
<th>Precision among days</th>
<th>Inter-batch precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>CA199</td>
<td>Lot 1</td>
<td>115.337</td>
<td>5.373</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>114.343</td>
<td>5.327</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>115.738</td>
<td>6.736</td>
</tr>
<tr>
<td></td>
<td>Lot 1</td>
<td>168.59</td>
<td>10.535</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>166.09</td>
<td>11.975</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>164.91</td>
<td>11.228</td>
</tr>
<tr>
<td>AFP</td>
<td>Lot 1</td>
<td>68.298</td>
<td>5.520</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>66.648</td>
<td>4.691</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>67.32</td>
<td>4.90</td>
</tr>
<tr>
<td>PGI</td>
<td>Lot 1</td>
<td>5.018</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>4.853</td>
<td>0.308</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>5.062</td>
<td>0.414</td>
</tr>
<tr>
<td>PGII</td>
<td>Lot 1</td>
<td>21.976</td>
<td>2.101</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>21.673</td>
<td>1.947</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>21.559</td>
<td>1.751</td>
</tr>
<tr>
<td>CEA</td>
<td>Lot 1</td>
<td>146.985</td>
<td>13.438</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>147.287</td>
<td>13.050</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>146.624</td>
<td>12.214</td>
</tr>
<tr>
<td>NSE</td>
<td>Lot 1</td>
<td>12.580</td>
<td>1.027</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>12.363</td>
<td>1.083</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>12.494</td>
<td>0.946</td>
</tr>
<tr>
<td>Pro-GRP</td>
<td>Lot 1</td>
<td>0.409</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>0.405</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>0.401</td>
<td>0.024</td>
</tr>
<tr>
<td>CA153</td>
<td>Lot 1</td>
<td>10.271</td>
<td>0.897</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>10.219</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>10.149</td>
<td>0.904</td>
</tr>
<tr>
<td>βHCG</td>
<td>Lot 1</td>
<td>8.519</td>
<td>0.749</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>8.608</td>
<td>0.715</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>8.747</td>
<td>0.636</td>
</tr>
<tr>
<td>CA724</td>
<td>Lot 1</td>
<td>19.369</td>
<td>1.686</td>
</tr>
<tr>
<td></td>
<td>Lot 2</td>
<td>19.691</td>
<td>1.776</td>
</tr>
<tr>
<td></td>
<td>Lot 3</td>
<td>19.346</td>
<td>1.398</td>
</tr>
</tbody>
</table>
Supplementary Table 5. Diagnostic models established from multiple biomarkers for various cancer prediction

<table>
<thead>
<tr>
<th>Gender</th>
<th>Biomarker*</th>
<th>Diagnostic model</th>
<th>AUC</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>PGI</td>
<td>$Y = \logit(P) = -1.518 - 0.007X_{PGI} + 0.001X_{CA199} - 0.002X_{CA125} + 0.005X_{CEA} - 0.396X_{pro-GRP} + 0.005X_{CA724}$</td>
<td>0.683</td>
<td>40.9</td>
<td>76.7</td>
</tr>
<tr>
<td></td>
<td>PGII</td>
<td>$Y = \logit(P) = -1.237 - 0.017X_{PGI} + 0.001X_{CA199} - 0.02X_{NSE} + 0.005X_{CEA} - 0.277X_{pro-GRP} + 0.005X_{CA724}$</td>
<td>0.690</td>
<td>39.5</td>
<td>71.4</td>
</tr>
<tr>
<td></td>
<td>pro-GRP</td>
<td>$Y = \logit(P) = -1.993 - 0.004X_{PGI} + 0.13X_{NSE} + 0.001X_{CA125} + 0.544X_{pro-GRP}$</td>
<td>0.719</td>
<td>21.0</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td>AFP</td>
<td>$Y = \logit(P) = -3.224 + 0.003X_{AFP}$</td>
<td>0.856</td>
<td>68.2</td>
<td>92.0</td>
</tr>
<tr>
<td></td>
<td>NSE</td>
<td>$Y = \logit(P) = -2.149 + 0.01X_{NSE} + 0.007X_{CEA} + 0.625X_{pro-GRP}$</td>
<td>0.785</td>
<td>62.4</td>
<td>56.4</td>
</tr>
<tr>
<td></td>
<td>CEA</td>
<td>$Y = \logit(P) = -1.68 - 0.005X_{PGI} + 0.001X_{CA199} - 0.023X_{NSE} + 0.005X_{CEA} + 0.003X_{CA724}$</td>
<td>0.706</td>
<td>67.9</td>
<td>67.2</td>
</tr>
<tr>
<td></td>
<td>CA125</td>
<td>$Y = \logit(P) = -0.567 - 0.007X_{PGI} + 0.013X_{NSE} + 0.002X_{CA199} + 0.018X_{NSE} + 0.012X_{CEA} + 0.436X_{pro-GRP} + 0.005X_{CA724}$</td>
<td>0.718</td>
<td>55.5</td>
<td>63.7</td>
</tr>
<tr>
<td></td>
<td>CA199</td>
<td>$Y = \logit(P) = -1.924 - 0.007X_{PGI} + 0.014X_{PGII} + 0.001X_{CA199} - 0.002X_{NSE} + 0.007X_{CEA} + 0.006X_{pro-GRP} + 0.005X_{CA724}$</td>
<td>0.693</td>
<td>42.6</td>
<td>74.9</td>
</tr>
<tr>
<td></td>
<td>CA724</td>
<td>$Y = \logit(P) = -1.816 - 0.005X_{PGI} + 0.001X_{CA199} - 0.371X_{pro-GRP} + 0.009X_{CA724}$</td>
<td>0.674</td>
<td>61.8</td>
<td>67.2</td>
</tr>
<tr>
<td></td>
<td>total-PSA</td>
<td>$Y = \logit(P) = -4.29 + 0.067X_{total-PSA}$</td>
<td>0.867</td>
<td>81.0</td>
<td>88.7</td>
</tr>
<tr>
<td></td>
<td>free-PSA</td>
<td>$Y = \logit(P) = -4.145 + 0.045X_{total-PSA}$</td>
<td>0.909</td>
<td>84.0</td>
<td>79.3</td>
</tr>
<tr>
<td>Female</td>
<td>PGI</td>
<td>$Y = \logit(P) = -1.385 - 0.005X_{PGI} + 0.001X_{CA199} - 0.057X_{NSE} + 0.004X_{CEA} - 0.626X_{pro-GRP} + 0.007X_{CA724}$</td>
<td>0.704</td>
<td>22.2</td>
<td>89.8</td>
</tr>
<tr>
<td></td>
<td>PGII</td>
<td>$Y = \logit(P) = -1.359 - 0.005X_{PGI} + 0.001X_{CA199} + 0.057X_{NSE} + 0.004X_{CEA} - 0.622X_{pro-GRP} + 0.007X_{CA724}$</td>
<td>0.702</td>
<td>55.1</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td>pro-GRP</td>
<td>$Y = \logit(P) = -2.392 + 0.009X_{CEA} + 0.863X_{pro-GRP} + 0.004X_{CA153}$</td>
<td>0.804</td>
<td>28.2</td>
<td>99.7</td>
</tr>
<tr>
<td></td>
<td>AFP</td>
<td>$Y = \logit(P) = -3.201 + 0.003X_{AFP}$</td>
<td>0.780</td>
<td>55.9</td>
<td>87.2</td>
</tr>
<tr>
<td></td>
<td>NSE</td>
<td>$Y = \logit(P) = -1.938 - 0.005X_{PGI} + 0.004X_{CEA} + 1.531X_{pro-GRP} + 0.006X_{CA153} + 0.005X_{CA724}$</td>
<td>0.721</td>
<td>44.6</td>
<td>71.6</td>
</tr>
<tr>
<td></td>
<td>CEA</td>
<td>$Y = \logit(P) = -2.005 - 0.005X_{PGI} + 0.002X_{CA199}$</td>
<td>0.660</td>
<td>67.3</td>
<td>69.0</td>
</tr>
<tr>
<td></td>
<td>CA125</td>
<td>$Y = \logit(P) = -0.629 - 0.005X_{PGI} + 0.001X_{CA199} + 0.004X_{AFP} + 0.032X_{CEA} + 0.961X_{pro-GRP} + 0.008X_{CA153} + 0.007X_{CA724}$</td>
<td>0.744</td>
<td>58.4</td>
<td>52.4</td>
</tr>
<tr>
<td></td>
<td>CA199</td>
<td>$Y = \logit(P) = -2.234 + 0.001X_{CA199} + 0.005X_{CA724}$</td>
<td>0.681</td>
<td>51.0</td>
<td>79.6</td>
</tr>
<tr>
<td></td>
<td>CA724</td>
<td>$Y = \logit(P) = -1.861 - 0.005X_{PGI} + 0.002X_{CA199} + 0.002X_{CA125} + 0.004X_{CEA} - 0.529X_{pro-GRP} - 0.013X_{CA153} + 0.015X_{CA724}$</td>
<td>0.764</td>
<td>47.4</td>
<td>80.1</td>
</tr>
<tr>
<td></td>
<td>βHCG</td>
<td>$Y = \logit(P) = -2.971 - 0.009X_{CA199} + 0.001X_{CA125} + 0.004X_{βHCG}$</td>
<td>0.701</td>
<td>33.3</td>
<td>81.1</td>
</tr>
<tr>
<td></td>
<td>CA153</td>
<td>$Y = \logit(P) = -3.184 + 0.009X_{CA153}$</td>
<td>0.649</td>
<td>47.9</td>
<td>86.2</td>
</tr>
</tbody>
</table>

Note: *Biomarker was separately detected by old method and was part of the protein chip results. AUC: Areas under the curve.

Supplementary Table 6. Diagnostic models established from multiple biomarkers for total disease or cancer prediction

<table>
<thead>
<tr>
<th>Gender</th>
<th>Function</th>
<th>Diagnostic model</th>
<th>AUC</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Disease prediction</td>
<td>$Y = \logit(P) = 8.325 + 0.028X_{PGI} + 0.057X_{AFP} + 0.455X_{total-PSA} + 2.641X_{pro-GRP} + 0.882X_{CEA} + 1.444X_{CA125} + 8.028X_{pro-GRP} + 0.057X_{CA724}$</td>
<td>0.981</td>
<td>91.2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Cancer prediction</td>
<td>$Y = \logit(P) = 0.197 + 0.005X_{PGI} + 0.01X_{AFP} + 0.001X_{CEA} - 0.008X_{CA125} + 0.144X_{pro-GRP} + 0.5X_{pro-GRP} + 0.003X_{CA724}$</td>
<td>0.691</td>
<td>46.1</td>
<td>83.1</td>
</tr>
<tr>
<td>Female</td>
<td>Disease prediction</td>
<td>$Y = \logit(P) = -0.384 + 0.374X_{NSE}$</td>
<td>0.836</td>
<td>71.8</td>
<td>95.2</td>
</tr>
<tr>
<td></td>
<td>Cancer prediction</td>
<td>$Y = \logit(P) = -0.614 + 0.006X_{PGI} + 0.002X_{CA199} + 0.003X_{AFP} - 0.03X_{CEA} + 1.307X_{pro-GRP} - 0.002X_{βHCG} + 0.016X_{CA153} + 0.009X_{CA724}$</td>
<td>0.753</td>
<td>51.6</td>
<td>89.2</td>
</tr>
</tbody>
</table>

AUC: Areas under the curve.