EBioMedicine xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Multiplexed Serum Biomarkers for the Detection of Lung Cancer

Shenglin Ma MD [a,c], Wenzhe Wang PhD [b,c], Bing Xia MD [a], Shirong Zhang MSc [a,c], Haining Yuan MSc [b,c], Hong Jiang MD [a], Wen Meng MD, MSc [a], Xiaoliang Zheng PhD [b,c], Xiaoju Wang PhD [b,c,*]

[a] Department of Radiation Oncology, Affiliated Hangzhou Hospital of Nanjing Medical University, Zhejiang Academy of Medical Sciences, Hangzhou, Zhejiang, PR China
[b] Center for Molecular Medicine, Zhejiang Academy of Medical Sciences, Hangzhou, Zhejiang, PR China
[c] Institute of Lung Cancer, Zhejiang Academy of Medical Sciences, Hangzhou, Zhejiang, PR China

ARTICLE INFO

Article history:
Received 4 May 2016
Received in revised form 1 August 2016
Accepted 12 August 2016
Available online xxxx

Keywords:
Lung cancer
Serum biomarkers
Risk factors
Early detection
LDCT

ABSTRACT

Currently, there is no available biomarker for lung cancer diagnosis. Here we recruited 844 lung cancer patients and 620 healthy participants from six hospitals. A total of four serum proteins was identified and subsequently assessed in the training and validation cohorts. The concentrations of four serum proteins were found to be significantly higher in lung cancer patients compared with healthy participants. The area under the curve (AUC) for the 4-biomarker were 0.86 in the training cohort, and 0.87 in the validation cohort. The classification improved to a corrected AUC of 0.90 and 0.89 respectively following addition of sex, age and smoking status. Similar results were observed for early-stage lung cancer. Remarkably, in a blinded test with a suspicious pulmonary nodule, the adjusted prediction model correctly discriminated the patients with 86.96% sensitivity and 98.25% specificity. These results demonstrated the 4-biomarker panel improved lung cancer prediction beyond that of known risk factors. Moreover, the biomarkers were valuable in differentiating benign nodules which will remain indolent from those that are likely to progress and therefore might serve as an adjuvant diagnosis tool for LDCT scanning.

1. Introduction

Lung cancer is continuously the leading cause of cancer-related deaths for both men and women worldwide (Cancer Facts & Figures 2015, American Cancer Society). The early detection of lung cancer presents an opportunity to dramatically reduce disease mortality. The overall 5-year relative survival rate was 17%; however, if the cancer is detected at stage Ia, the 5-year survival often exceeds 80% (Mahadevia et al., 2003). Currently, the National Lung Screening Trial (NLST) using low dose chest computed tomography (LDCT) in high-risk individuals demonstrates that a 20% reduction in lung cancer-specific mortality and a 6.7% reduction in all-cause mortality can be achieved (International Early Lung Cancer Action Program I et al., 2006; Bach et al., 2007; Henschke et al., 1999; Sone et al., 1998). However, there is a significant chance of a false-positive result for CT scan, which may require additional clinical testing, even an invasive procedure to specify an abnormality (Welch et al., 2007; Wilson et al., 2008; Swensen et al., 2005). Such finding greatly exacerbates the high cost of the technology and leads to unnecessary patient anxiety and surveillance. Therefore, a non-invasive test with a high specificity for distinguishing the indolent disease from lung cancer patients is highly demanded.

Serum biomarkers could be used as an invasive, cost-effective way to differentiate lung cancer patients. Several serum tumor markers have been studied extensively, such as carcinoembryonic antigen (CEA), serum cytokeratin 19 fragments, and pro-gastrin-releasing peptide; however, none has been demonstrated to provide clinical utility, mainly because of the poor reproducibility and lack of sufficient sensitivity and specificity (Buccheri et al., 2003; Pastor et al., 1997).

Recently, we described a non-invasive diagnostic system on Luminex xMAP platform to detect serum autoantibodies for diagnosis of lung cancer (Jia et al., 2014). Given the biological properties of cancer as a systemic disease, we predict that a combination of cancer associated serum proteins and autoantibody can be used to achieve superior levels of sensitivity and specificity. Here, we identified a diverse set of circulating proteins in the sera of patients with lung cancer and designed a large-scale, multicenter validation study to evaluate their utility in distinguishing lung cancer patients from matched healthy controls, with the goal of using these biomarkers to aid clinicians in making case management decisions.

2. Methods

2.1. Patients Population

The recruitment of patient with lung cancer was initiated from May 2009 and the discovery phase of the study included the patients collected until September 2009 from Hangzhou First People's Hospital,

* Corresponding author at: Center for Molecular Medicine, 182 Tianmushan Road, Hangzhou, Zhejiang 310012, PR China.
E-mail address: wangxj@zjams.com.cn (X. Wang).

http://dx.doi.org/10.1016/j.ebiom.2016.08.018

Zhejiang, China. The training cohort included the lung cancer patients collected between February 2011 and March 2012 from 3 hospitals, namely Hangzhou First People's Hospital, Hangzhou Cancer Hospital and Zhejiang Cancer Hospital, Zhejiang, China. The validation lung cancer patients were recruited from September 2012 to December 2013 from 6 hospitals, including the 3 hospitals mentioned in the training cohort, and the First Hospital of Jiaxing, the First Affiliated Hospital of Wenzhou Medical University, Shaoxing People's Hospital, Zhejiang, China.

Approval for the study was obtained from the institutional ethics review committee. All patients were provided written informed consent, according to the committees' regulations.

2.2. Patient and Serum Collection

The patients included in this study were all consecutive patients, and details regarding patient inclusion criteria and collection protocols have been previously reported (Jia et al., 2014). All cases used in this study were confirmed to be primary lung cancer by pathology review.

Healthy participants were recruited from the eligible blood donors with no evidence of pulmonary disorders of any type and were approximately age and sex-matched to the cancer cohorts presented. Patients who had a history of other solid tumors were excluded from the study.

2.3. Cancer Biomarker Screening and Detection

In discovery phase, samples were analyzed using the Bio-Plex Pro Biomarker Assays (Bio-Rad Laboratories), according to the manufacturer's instruction. CEA concentrations were measured with commercial ELISA (Roche), according to the manufacturer's recommendations. All measurements were done in duplicate.

2.4. Autoantigen Coupling to Luminex Microsphere and Detection of Autoantibody in Serum

The cDNA for NY-ESO-1 was cloned into the Flexi vector with a HQ-tag at N-terminus (Promega, USA) as previously described (Jia et al., 2014). The recombinant protein was expressed and purified using Ni-NTA agarose (Life Technologies) according to manufacturers' protocols. The purity was determined as >95% by SDS-PAGE and Coomassie staining.

A total of 25 μg of purified protein was conjugated to microspheres by following the manufacture's protocol (Luminex, Austin). The microspheres conjugated with recombinant protein were aliquoted into a 96-well plate, and autoantibody in serum was detected by a Bio-Plex 200 System as previously described (Jia et al., 2014). All values reported are the raw median fluorescence intensities (MFI).

2.5. Calibrator for Autoantibody Measurement

Autoantibody in serum was detected by a Bio-Plex 200 System as previously described (Jia et al., 2014). Currently, there are no calibration standards for assays to measure cancer autoantibodies. Therefore, a modified calibration system was adopted based on a previous publication (Murray et al., 2010). Briefly, serum samples from lung cancer patients were screened for the autoantibody against NY-ESO-1 as described above. The positive sera were then further confirmed by western blot. For each serum sample, a calibration curve of background-corrected MFI versus log dilution was constructed to which a four-parameter logistic model plot was fitted. The MFI value for each unknown sample was then converted to a calibrated reference unit (RU) using the calibration curve. A calibration curve was prepared at the beginning of every assay run.

2.6. Statistical Analysis

In the discovery phase, the differences of the circulating concentrations of each biomarker between two groups were evaluated by the Mann-Whitney *U* test (continuous variables and nonparametric analyses). R version 3.0.1, 2-sided tests and a significance level of 0.05 were used. If not stated otherwise, we considered all patients with lung cancer as a single group regardless of stage.

For the training and validation cohorts, statistical analyses were carried out with MedCalc (version 15.8). Receiver operating characteristics (ROC) curves were used to quantify the biomarker performance by means of sensitivity, specificity, area under the curve (AUC) as well as corresponding 95% confidence intervals. To test the diagnostic accuracy of the panel of biomarkers, we estimated functions of the combined markers by logistic regression with or without adjustment for known risk factors for lung cancer (age, sex and smoking status), and the predictive probabilities were used as one marker and subjected to ROC analysis. We investigated the optimum cutoff value for diagnosis by maximizing the sum of sensitivity and specificity and minimizing the overall error (square root of the sum $[1\text{-sensitivity}]^2 + [1\text{-specificity}]^2$), and by minimizing the distance of the cutoff value to the top-left corner of the ROC curve. We also reported adjusted p-values corrected for multiple testing using the Benjamini-Holm method to control for the false positive error rate.

The correlation between the biomarkers in serum and clinicopathological characteristics was analyzed with Fisher's exact test. We took $p < 0.05$ (two sided) to be significant.

2.7. Prediction of Blinded Patients

The serum samples were collected pre-surgery from Hangzhou First People's Hospital, Zhejiang, China. The biomarkers were measured as described above, and the blinded patients were predicted as cancer or non-cancer by the BRB-Array Tools package (version 3.6) available at http://linus.nci.nih.gov/BRB-ArrayTools.html. Briefly, a predictor model was created using the 4-biomarker panel after adjustment for age, sex and smoking status in the training set, then subsequently tested in the validation cohort. A log base 2 transformation was applied to the raw data. Each sample's value was multiplied by the corresponding coefficients derived from univariate logistic regressions on the training set with cancer/non-cancer as a binary response variable, and then the values were totaled. The adjusted index scores were then assessed by the ROC curve, which provided a pure index of a test's accuracy by plotting the sensitivity against 1-specificity for each result value of the test. We computed the misclassification error of the models using leave-one-out cross-validation (LOOCV) method. For each LOOCV training set, the entire model-building process was repeated, including the biomarker selection process. The class labels were randomly permuted (100 permutations), and the entire LOOCV process was repeated. The significance level is the proportion of the random permutations that gave a cross-validated error rate no greater than the cross-validated error rate obtained with the real data. Each clinical sample was predicted to cancer or non-cancer group by the model. We computed a statistical significance level ($p < 0.05$) for each biomarker.

3. Results

A total of 844 patients with lung cancer were included in this study, 40 in the discovery cohort, 543 in the training cohort and 261 in the validation cohort (Fig. 1). The healthy controls included 620 healthy participants. There is no significant difference in term of age and sex in both case and control groups, however, more current smokers in cases than controls. Clinicopathological characteristics of the study participants are summarized in Table 1. We also recruited 70 patients with various benign lung diseases, and 80 blinded patients with suspicious pulmonary nodule detected by LDCT. All these high-risk patients were either

Please cite this article as: Ma, S., et al., Multiplexed Serum Biomarkers for the Detection of Lung Cancer, EBioMedicine (2016), http://dx.doi.org/10.1016/j.ebiom.2016.08.018