Blood Test for Early Detection of Lung Cancer

Martin Stengel,1 Harvey Pass,2 William Rom,2 Sudeep Kumar,1 Srikanth Vaithilingam,1 Anahit Aghvanyan,1 Daisy Roy,1 Gelu Dobrescu,1 Ramou Sivakamasundari,1 Eli N. Gezer,1 and Jacob N. Wohlstätter1

1Meso Scale Diagnostics, LLC, Rockville, Maryland, USA. 2New York University, School of Medicine, New York, NY, USA

Abstract

Lung cancer is the largest single cause of death from cancer worldwide. Even though lung cancer often can be treated successfully when detected early, approximately 90% of patients diagnosed with lung cancer ultimately succumb to the disease. Screening with low dose computed tomography (CT) can reduce mortality, but the positive predictive value of this test is low, leading to a large number of suspicious but ultimately non-malignant results that nevertheless require follow-up. Our objective was to develop a simple blood test to risk-stratify patients at high risk of lung cancer. We developed multiplexed, serum/plasma immunoassay panels to measure more than 40 lung cancer-related biomarkers using a 96-well, 7-spot format and electrochemiluminescence detection. Due to the high sensitivity of MSD's MULTI-ARRAY® technology, these panels were run with diluted serum or plasma, bringing the total sample volume required to run all 40 assays down to approximately 40 µL per replicate. This enabled us to measure all markers simultaneously in precious, high-quality serum and EDTA plasma samples. We used samples from early-stage lung cancer patients (drawn before lung cancer surgery) and from a lung-cancer screening cohort of age-matched heavy smokers who did not have lung cancer at the time of the blood draw. In a training set of 300 samples, 12 serum and 8 plasma markers had areas under an ROC curve (AUC) of 0.7 or higher. We used a logistic regression model with 100x cross-validation to develop a multi-marker panel. One serum panel (Flt-3L, EGFR, MMP-3, and NME-2) and one plasma panel (Flt-3L, cytokeratin-19, MMP-3, Flt-1, KGF, and PFG) were selected and tested using approximately 250 additional samples from the same cohort. For the serum panel, the ROC area dropped to 0.95 (vs. 0.95 for the training set) for the plasma panel, the ROC area dropped to 0.81 (vs. 0.80). The ROC area of 0.85 for the serum panel with clinical sensitivity and specificity of 81% and 84%, respectively, and the ROC area of 0.81 for the plasma panel (with clinical sensitivity and specificity of 76% and 81%, respectively), are expected to be clinically useful despite this ROC area decrease.

Analysis of the combined training and test sets with 100x cross-validation resulted in a 4-marker serum panel (Flt-3L, EGFR, MMP-3, and NME-2) with an ROC area of 0.91 and clinical sensitivity and specificity of 88% and 82%, respectively, and a 5-marker plasma panel (Flt-3L, cytokeratin-19, Flt-1, KGF, and PFG) with an ROC area of 0.91 and clinical sensitivity and specificity of 84% and 87%, respectively. Using MULTI-ARRAY® technology and high quality clinical samples, we were able to identify promising biomarker panels for early detection of lung cancer in high-risk individuals.

Methods

MSD's electrochemiluminescence detection technology uses SULFO-TAG™ labels that emit light upon electrochemical stimulation initiated at the electrode surfaces of MULTI-ARRAY® and MULTI-SPOT® multiplexers.

Electrochemiluminescence Technology

- Minimal non-specific background and strong responses to analyze yield high signal-to-background ratios.
- The stimulation mechanism (electricity) is decoupled from the response (light signal), minimizing matrix interference.
- Only labels bound near the electrode surface are excited, enabling non-washed assays.
- Labels are stable, non-radioactive, and directly conjugated to biological molecules.
- Emission at >200 nm eliminates problems with color quenching.
- Multiple rounds of label excitation and emission enhance light levels and improve sensitivity.
- Carbon electrode surface has 10X greater binding capacity than polystyrene wells.

Analytical Protocol

1. Add 200 µL MSD® Blocker A to each well. Incubate for 30 min at room temperature (RT).
2. Wash with PBS-T. Add 25 µL of assay diluent and 25 µL of diluted sample or calibrator. Incubate for 2 hours at RT with shaking. (Note: For the last three panels in the bottom left table [Ca 125, ..., uPA], add antibody. Incubate for 1 hour at RT.
3. Wash with PBS-T. Add 150 µL of Read Buffer T. Read on MSD® reader (Note: For the 3rd panel [Ca 15, ..., CPM], use Read Buffer P).

One 150 µL serum and one 150 µL EDTA-plasma aliquot was available from each of the 400 individual (200 cases, 200 matched controls). This was sufficient to test all 47 biomarkers in duplicates. Each plate had an approximately equal number of cases and controls, and samples were run and analyzed blinded. Serum and plasma samples were selected and analyzed independently.

Data Analysis

- Use logistic regression to model combinations of assays

ROC Analysis of Training and Validation Set

<table>
<thead>
<tr>
<th>Marker</th>
<th>Serum Panel</th>
<th>Plasma Panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flt-3L</td>
<td>0.95</td>
<td>0.81</td>
</tr>
<tr>
<td>EGFR</td>
<td>0.85</td>
<td>0.81</td>
</tr>
<tr>
<td>MMP-3</td>
<td>0.85</td>
<td>0.81</td>
</tr>
<tr>
<td>NME-2</td>
<td>0.85</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Conclusion

- MULTI-ARRAY® panels containing 45 lung cancer-related markers were developed and analytically validated.
- The high sensitivity, dynamic range, and multiplexing of MSD assays allowed all 45 markers to be measured in a 40 µL sample volume.
- Approximately 550 serum and EDTA-plasma samples from early stage lung cancer patients and from heavy smokers who do not have lung cancer were run on these panels in several batches over a period of nine months.
- An initial training set of 150 serum samples and ~150 EDTA-plasma samples was used to develop a serum algorithm and a plasma algorithm to separate cases from controls. These algorithms were then validated using an independent validation set (~250 samples).
- The ROC area for the serum validation set was 0.85, and for the plasma validation set 0.81.
- Using the entire data set for both training and validation (including cross-validation), the ROC area for the best 4-marker serum panel (containing Flt-3L, EGFR, MMP-3, and NME-2) was 0.91.
- Using the entire data set for both training and validation (including cross-validation), the ROC area for the best 5-marker plasma panel (containing Flt-3L, Cytokeratin-19, Flt-1, KGF, and HGF) was 0.91.

This project has been funded in whole or in part with Federal funds from the National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN23120100010C.