How Robust Are Your Archives?

Harvey I. Pass MD
NYU Langone Health
New York, NY
Disclosures

• **Grants:** NCI EDRN, TCGA, DOD, Technion, CDC, CPTAC

• **Philanthropy:** Rosenwald Foundation, Stephen A. Banner Lung Cancer Foundation, Belluck and Fox, Levi, Phillips and Konigsberg, Simmons Foundation, Baron and Budd, multiple patients who shall remain anonymous

• **Royalties:** NCI cell lines

• **Industry:** Fujirebio, 20/20 Gene, MesoScale Diagnostics, Cynvezio, Source MDX, Celera, OPKO, SomaLogic, Genentech, Integrated Diagnostics, Transgenomics, Rosetta Genomics, Calithera, Pinpoint Genomics, Nanostring, Cizzle, EmeraldLogic, HTG, Freenome, Biodesix

• **Patents** (no money): Osteopontin for diagnosis and prognosis of MPM; fibulin 3 for diagnosis and prognosis of MPM; mir-29c* for prognosis of MPM; mir-31 for diagnosis of MPM, HMGB1 for diagnosis of MPM
Surgeons Who Want To Do Translational Science

• Usually are very focused with regard to what they are studying
• Have a lot of resources
• Know their limitations but overcome limitations by collaboration
• Don’t stop operating and in fact recruit other surgeons to collaborate
• **Have an organized, prospective collection system which allows them complete control of their own specimens with parallel prospective recording of diagnostic and prognostic endpoints.**
• **They are proud specimen hoarders, who, frankly, want total control of their science and can't stand waiting for answers.**
• **But....they must be willing to share their resources and realize that 99% of the people who may use these resources are smarter than them.**
The NYU Lung Cancer Thoracic Surgery Archives

• Purpose
  • Patient data prospectively with matching specimens which could potentially improve the diagnosis and prognosis of resectable lung cancers.

• 2006-present
  • NLST: False positive prevention for defining the CT nodule, need for early detection non-imaged based strategies, need for ways to define aggressive vs indolent early stage lung cancers
  • IASLC Staging System x 2, IASLC Adenocarcinoma Lung Cancer Classification, mutation detection besides tissue biopsy, surrogate profiles for prognostication of early stage lung cancer, recognition of immunooncology for therapy prediction

Many questions which, frankly, were unanswerable without robust databases and with matching specimen archives the potential for novel solutions to these questions is limited only by novel ways to use the specimens.
Logistics

• Clinic
  • consent
  • notify lab
  • prepare baggies and liquid nitrogen for the week

• OR Day
  • approach missed patients in presurgical
  • Research Assistant magically appears in surgeons room
  • Blood from patient at induction
  • Specimen resection
    • Surgeon takes tumor and normal lung at table
    • Places in nunc vials
    • Immediately placed in Liquid Nitrogen

• Serum, plasma, PBMC, buffy coat, effusion cell/liquid, paxgene processed and stored at -80°C within 4 hours barcoded in LabVantage in Thoracic Surgery Lab

• REDCAP list of all specimens, demographics, pathology

• Weekly lists of returning patients on study for followup serum and plasma.
  • Longitudinal followup 3 months x 2 years, 6 months x 1 year, then yearly
# Overview

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number Individuals</th>
<th>Plasma</th>
<th>Serum</th>
<th>Buffy Coat</th>
<th>PBMC</th>
<th>Tumor/Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Risk Smokers/Ex Smokers</td>
<td>1100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign Lung Pathology resected in OR</td>
<td>305</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total NSCLC</td>
<td>1200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>793</td>
</tr>
<tr>
<td>Stage I lung cancers</td>
<td>800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>560</td>
</tr>
<tr>
<td>Stage II lung cancers</td>
<td>132</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesotheliomas (NYU)</td>
<td>147</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115</td>
</tr>
</tbody>
</table>
NSCLC Survival by Stage NYU Archives

Number at risk
Group: 1
443 370 294 235 166 120 83 47 19 0
Group: 2
99 72 53 43 35 24 13 5 1 0

NSCLC Survival by Stage Adenocarcinoma

Number at risk
Group: 1
356 301 249 202 145 104 73 41 16 0
Group: 2
74 53 37 30 24 18 10 5 1 0

Are These Archives Representative?
Problems:
1. Previous Pulmonary resections
2. Not lung cancer specific death
3. Less than 6 lymph nodes sampled
<table>
<thead>
<tr>
<th></th>
<th>Males (94)</th>
<th>Females (210)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>71±1</td>
<td>69±1</td>
</tr>
<tr>
<td><strong>Pack-Years</strong></td>
<td>37±3</td>
<td>26±2</td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>1.9±0.9</td>
<td>1.9±0.1</td>
</tr>
<tr>
<td><strong>LVI</strong></td>
<td>12 (13%)</td>
<td>28 (13%)</td>
</tr>
<tr>
<td><strong>Pleural Invasion</strong></td>
<td>13 (14%)</td>
<td>25 (12%)</td>
</tr>
<tr>
<td><strong>Positive Nodes</strong></td>
<td>5 (5%)</td>
<td>13 (6%)</td>
</tr>
<tr>
<td><strong>Resection Type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wedge</td>
<td>11 (12%)</td>
<td>26 (12%)</td>
</tr>
<tr>
<td>Segment</td>
<td>18 (19%)</td>
<td>49 (23%)</td>
</tr>
<tr>
<td>Lobectomy</td>
<td>65 (69%)</td>
<td>135 (65%)</td>
</tr>
</tbody>
</table>
IASLC Staging 8: Sort of…

**cStage I NSCLC Adenocarcinoma
Survival by Size**

<table>
<thead>
<tr>
<th>Time (days)</th>
<th>Number at risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 0-1</td>
</tr>
<tr>
<td></td>
<td>26 26 26 25 22 18 16 12 11 7 2 0</td>
</tr>
<tr>
<td></td>
<td>Group 1-2</td>
</tr>
<tr>
<td></td>
<td>147 140 135 109 85 67 46 33 24 15 4 0</td>
</tr>
<tr>
<td></td>
<td>Group 2-3</td>
</tr>
<tr>
<td></td>
<td>77 77 69 52 39 30 18 11 6 3 1 0</td>
</tr>
<tr>
<td></td>
<td>Group &gt;3</td>
</tr>
<tr>
<td></td>
<td>35 35 33 24 18 12 5 1 0 0 0 0 0</td>
</tr>
</tbody>
</table>

**cStage I NSCLC Adenocarcinoma
Including Node Positive on Path**

<table>
<thead>
<tr>
<th>Time (days)</th>
<th>Number at risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 0-1</td>
</tr>
<tr>
<td></td>
<td>27 27 27 26 23 19 17 13 12 8 3 0</td>
</tr>
<tr>
<td></td>
<td>Group 1-2</td>
</tr>
<tr>
<td></td>
<td>150 143 138 112 88 70 49 34 25 16 4 0</td>
</tr>
<tr>
<td></td>
<td>Group 2-3</td>
</tr>
<tr>
<td></td>
<td>87 86 78 59 44 34 21 14 7 4 1 0</td>
</tr>
<tr>
<td></td>
<td>Group &gt;3</td>
</tr>
<tr>
<td></td>
<td>40 39 35 26 20 14 6 2 1 0 0 0 0</td>
</tr>
</tbody>
</table>
IASLC Adenocarcinoma Histologic Categorization: Sort of...
Influence of Resection Type: Sort of….
### Multivariate Analysis Results

<table>
<thead>
<tr>
<th></th>
<th>Multivariate p Value</th>
<th>HR</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.0356</td>
<td>1.0463</td>
<td>1.0030 to 1.0913</td>
</tr>
<tr>
<td>Mayo Classification</td>
<td>0.0004</td>
<td>4.4172</td>
<td>1.9440 to 10.0370</td>
</tr>
<tr>
<td>Gender</td>
<td>0.0074</td>
<td>0.3586</td>
<td>0.1693 to 0.7596</td>
</tr>
<tr>
<td>LVI</td>
<td>0.0281</td>
<td>2.4609</td>
<td>1.1016 to 5.4976</td>
</tr>
<tr>
<td>Positive nodes</td>
<td>0.0683</td>
<td>2.5293</td>
<td>0.9326 to 6.8600</td>
</tr>
<tr>
<td>Pack-years</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleural invasion</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resection type</td>
<td>ns</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Survival Analysis

- **Group 1:**
  - Number at risk:
    - Group 1: 137, 137, 130, 106, 87, 71, 47, 35, 28, 16, 5, 0

- **Group 2:**
  - Number at risk:
    - Group 2: 137, 132, 122, 96, 69, 50, 33, 23, 15, 10, 3, 0

- **P < 0.0001**
  - HR: 8.37 (4.27 – 16.78)
Key Questions for Clinical Stage I Lung Cancer

- Predicting which patients *WILL* develop a lung cancer
- Diagnosing whether a CT detected nodule is lung cancer or not without invasive means
- Prognostication whether the non invasively diagnosed lung cancer
  - Will be indolent or aggressive, i.e. do you have to operate at all?
  - Is one that has a high probability of recurrence
- What Platforms?
  - Proteomics
    - MRM
    - SomaMers
    - Autoantibody
  - microRNA
    - Exosomal serum
    - NanoString Plasma
    - HTG Serum
    - Nanostring Buffy Coat Immunoexpression
Conclusions

• Before starting any type of prognostic or diagnostic biomarker discovery, you need to confirm that your cohorts approximate clinical demographics and intermediate endpoints described in the literature.

• Standard SOPs must be consistently used.
Take Home

• Archives and Databases need a lot of care and feeding.
• Archives and Databases CANNOT be siloed
• Before starting any type of prognostic or diagnostic biomarker discovery, you need to confirm that your cohorts approximate clinical demographics and intermediate endpoints described in the literature.
• Standard SOPs must be consistently used.
• Push surgeons to help you get what you need (in an organized fashion)
• Support BOOSTER!!!!!!
Thanks to...

- **NYU Thoracic Lab**
  - Chandra Goparaju PhD
  - Jessica Donington MD
  - Alla Ivanova, PhD
  - Sergei Ivanov, PhD
  - Mario Cioce MD, PhD
  - Ricardo Cazzoli, PhD
  - Claudia Canino, MS
  - Ryan Harrington BS
  - Brady Kwong, BA
  - Amanda Beck BS, MD
  - Joe Levin BS, MD
  - Nathalie Hirsch BA, MD
  - Yvonne Owusu-Osprung, BA
  - Jordan Preiss, BA
  - Ki Prokrym BA
  - Allen Vaynblatt BA
  - Audrey Sorensen RN
  - Jennifer Hull, RN
  - Christopher Reid, BA

- **NYU Biostatistics**
  - Judy Goldberg MS

- **NYU Pulmonary Medicine**
  - William Rom MD, PhD

- **NYU Genomics CORE**
  - Adriana Heguy, PhD

- **NYU Histopathology CORE**
  - Cynthia Loomis MD, PhD

- **NYU Biostatistics**
  - Judy Goldberg MS

- **NYU Pulmonary Medicine**
  - William Rom MD, PhD

- **NYU Genomics CORE**
  - Adriana Heguy, PhD

- **NYU Histopathology CORE**
  - Cynthia Loomis MD, PhD

- **Mt. Sinai Selikoff Foundation**
  - Stephen Levin MD

- **DMCC**
  - Mark Thornquist PhD

- **Carbone Laboratory, University of Hawaii**
  - Michele Carbone MD, PhD
  - Haining Yang, PhD

- **Antoine Laboratory, University of Liverpool**
  - Dan Antoine PhD

- **UCLA Proteomics CORE**
  - Shreya Kanodiya, PhD

- **University of Toronto**
  - Ming-Sound Tsao MD
  - Geoffrey Liu MD
  - Marc De Perrot MD

- **NCI Early Detection Research Network**
- **CDC National Mesothelioma Virtual Bank**
- **NCI TCGA Mesothelioma**
- **Department of Defense CDMRP**
- Technion
  - Israel Vladovsky PhD
- **SomaLogic**
  - Rachel Ostroff PhD
- **Nanostring Statistical Section**
  - Wenjie Wu PhD