Immune Approach To Primary Graft Dysfunction
Exosomes, Proteomics and Machine Learning
Barry Fine MD PhD
Columbia University Medical Center
Disclosure

Funding: Department of Defense, the NIH and the Gerstner Foundation
Learning Objectives

1. Delineate the role of proteomics in biomarker discovery as it relates to primary graft failure

2. Understand how we can use machine learning to analyze complex data sets
Nick Giangreco
Three Themes

Exosomes

PGD

Machine Learning

Proteomics
Primary Graft Dysfunction

Graft dysfunction within 24 hours of transplant without a discernible cause such as rejection, pHTN, infection, etc...

Table 6: Definition of Severity Scale for Primary Graft Dysfunction (PGD)

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild PGD-LV</td>
<td>One of the following criteria must be met:</td>
</tr>
<tr>
<td>Moderate PGD-LV</td>
<td>Must meet one criterion from I and another criterion from II:</td>
</tr>
<tr>
<td>Severe PGD-LV</td>
<td>Diagnosis requires either both I and II, or III alone:</td>
</tr>
</tbody>
</table>

- UEF ≤ 40% by echocardiography, or Hemodynamics with RAP > 15 mm Hg, PCWP > 20 mm Hg, CI < 2.0 L/min/m² (lasting more than 1 hour) requiring low-dose inotropes
- Hemodynamic compromise with RAP > 15 mm Hg, PCWP > 20 mm Hg, CI < 2.0 L/min/m², hypotension with RAP < 70 mm Hg (lasting more than 1 hour)
- High-dose inotropes—Inotrope score > 10 † or Newly placed IABP (regardless of inotropes)
- Dependence on left or biventricular mechanical support including ECMO, LVAD, BiVAD, or percutaneous LVAD. Excludes requirement for IABP.
- Hemodynamics with RAP > 15 mm Hg, PCWP < 15 mm Hg, CI < 2.0 L/min/m², TPG < 15 mm Hg and/or pulmonary artery systolic pressure < 50 mm Hg, or Need for RVAD

† Inotrope score = dopamine (×1) + dobutamine (×1) + amrinone (×1) + milrinone (×15) + epinephrine (×100) + norepinephrine (×1000) with each drug dosed in μg/kg/min.

OPTN 2016
Primary Graft Dysfunction – Clinical Impact

A Kaplan Meier Survival Analysis: Impact of Severe PGD on Post-HT Outcomes in BTT Patients

A Receiver Operating Characteristic Analysis

B Percent of BTT Patients with Severe PGD Stratified by Number of Pre-Transplant Risk Factors

Truby et al JHLT 2018
Primary Graft Dysfunction - Etiology

Donor  Graft  Receptient

Acute Myocardial Dysfunction
Etiology? Opportunity?
Primary Graft Dysfunction – Precision Medicine

“an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person”

Donor

-NIH
Primary Graft Dysfunction – Precision Medicine

Donor

“an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person”

- NIH

Hostile Donor?

Circulating factors in the blood of transplant recipients impede cardiomyocyte function at the cellular level, leading to overall organ dysfunction.
Primary Graft Dysfunction – Hostile Donor
Exosomes

Characterize the “Hostile Milieu” Hypothesis

- Small 30-120nm cell derived vesicles that are likely present in all body fluids
- Released by cells as MVB merge with the cytoplasmic membrane
- Complex cargo including miRNA, mRNA, proteins and specialized lipids
- Delivered to recipient cells through endocytosis and have
- Burgeoning understanding of the role of exosomes in disease
Exosomes

- Can be rapidly isolated using immuno-affinity methods or ultrafiltration from accessible fluids (blood/urine)
- Secreted in both normal and pathological states
- Exosomal miRNA have been shown to be clinically relevant biomarkers for multiple cancers
- Exosomal proteins have been found to be potentially useful biomarkers in cancer, neurological disease and kidney disease
- Importantly, cargo is very **stable** compared to proteins and RNA in serum

<table>
<thead>
<tr>
<th>Protein</th>
<th>Disease Biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetulin-A</td>
<td>AKI</td>
</tr>
<tr>
<td>EGFRvIII</td>
<td>Glioblastoma</td>
</tr>
<tr>
<td>Phospho-Tau</td>
<td>Alzheimers</td>
</tr>
<tr>
<td>α-synuclein</td>
<td>parkinsons</td>
</tr>
</tbody>
</table>
Proteomics on Pre-Transplant Serum Exosomes

1. Generate exosome lysate
2. Denature, Reduce, Alkylate & Trypsin Digest
3. TMT labeling for quantification
4. Combine & Perform Peptide fractionation
5. Statistical Analysis
4. MS Data Acquisition
4. MS Quantification
Thermo Fusion Tribrid MS
Workflow

Data Generation

Proteome Discoverer

peptide identification & protein estimation

Columbia University

Cedars-Sinai

Workflow Evaluation cohorts
N=43 N=36

Follow the Clinical Laboratory Improvement Amendments... learning model selected PGD predictive proteins

Patient Data
Exosome
peptide identification & protein estimation
cohorts

https://github.com/ngiangre/cohorts

cohorts is:
• Modular
  • Each cohort is its own data structure
• Customizable
  • Each cohort has attributes for individual and integrative cohort analyses
• Flexible
  • Can use built in functions and use custom code leveraging the cohort attributes
There are two questions we can try to answer with proteomics

1. **Mechanistic**: Do the proteins tell us something about PGD. Is there a pathway discernible in differential protein analysis?

2. **Classification**: Can we use a protein or a set of proteins to help us predict PGD. Are there “features” can correctly classify transplant candidates at high risk versus those at low risk of PGD
Proteomics

Data distribution

<table>
<thead>
<tr>
<th></th>
<th># Data Files</th>
<th># Proteins</th>
<th># Replicates</th>
<th># Samples</th>
<th># Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columbia</td>
<td>2</td>
<td>1204</td>
<td>64</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>Cedar-Sinai</td>
<td>2</td>
<td>729</td>
<td>150</td>
<td>44</td>
<td>7</td>
</tr>
<tr>
<td>Paris</td>
<td>2</td>
<td>934</td>
<td>108</td>
<td>29</td>
<td>8</td>
</tr>
</tbody>
</table>
Prospective Cohort - Columbia

PGD Patients Have Different Exosome Proteomics Prior to Transplant
Prospective Cohort - Columbia

PGD Patients Have Different Exosome Proteomics Prior to Transplant
Retrospective Cohort – Cedars and Pitie Salpetriere

Low Correlation
Small number of differential proteins
PGD did not separate out
Machine Learning: Building mathematical models using training data in order to make predictions or decisions on test data without explicitly programmed to perform the task. Important to this is progressive improvement in that task with more data and training.
Classification Machine Learning

Different Algorithms

Multidimensional Feature Space

Allows identification of boundaries between two states
Classification Machine Learning

Different Algorithms

Multidimensional Feature Space

Allows identification of boundaries between two states
Classification Machine Learning

Different Algorithms

Multidimensional Feature Space

Allows identification of boundaries between two states
Classification Machine Learning

Different Algorithms

Multidimensional Feature Space

Allows identification of boundaries between two states

![Diagram showing classification in a multidimensional feature space with "Test" and "Training" sets.]
Classification Machine Learning

TPR | AUROC | FPR
---|---|---

Feature 1

Feature 3

"Training"

"Test"

"Training"
PGD—Prediction Using Logistic Regression

<table>
<thead>
<tr>
<th>Feature</th>
<th>AUROC</th>
<th>TPR</th>
<th>TNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Two Proteins</td>
<td>0.803</td>
<td>0.57</td>
<td>0.85</td>
</tr>
</tbody>
</table>
**PGD—Prediction Using Support Vector Machines**

- **Protein Expression**
- **Feature**
  - Best 2 Protein: AUROC 0.805, TPR 0.681, TNR 0.938
PGD—Prediction Using Random Forest

Feature

<table>
<thead>
<tr>
<th>Feature</th>
<th>AUROC</th>
<th>TPR</th>
<th>TNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best 2 Protein</td>
<td>0.807</td>
<td>0.660</td>
<td>0.875</td>
</tr>
</tbody>
</table>
PGD– Gradient Boosting

<table>
<thead>
<tr>
<th>Feature</th>
<th>AUROC</th>
<th>TPR</th>
<th>TNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best 2 Protein</td>
<td>0.829</td>
<td>0.72</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Best boundary may not be clinical
Bootstrapping with Recursive Feature Extraction

Constraint to One Protein

Constraint to Two Protein

More Dimensions $\rightarrow$ Better Prediction
Overall Prediction Performance

Feature Type:
- all three clinical variables
- all 265 proteins
- top 3 2-protein models

Model:
- Logistic Regression
- Random Forest
- Support Vector Machine
- Gradient Boosting Classifier

KLKB1 ELISA

p<0.0001
Acknowledgements

Biomedical Engineering
Gordana Vunjak-Novakovic PhD
Kacev Ronaldson PhD
Diogo Teles
Michael Kim

Columbia Cardiology – CHF/TX
Mary Jane Farr MD
Paolo Colombo MD
Hiroo Takayama MD PhD
Koji Takeda MD
Susan Restaino MD

Columbia Center for Translational Immunology
Emmanuel Zorn PhD
Poulomi Roy
Kortney Rogers

Columbia University Shared Proteomics Facility
Emily Chen PhD
Rajesh Soni PhD

CU Department of Bioinformatics
Nicholas Tatonetti PhD
Nicholas Giangreco

Hopital Pitie Salpetriere
Guillaume Lebreton MD
Pascal LePrince MD
Jean Luc Taupin MD

Cedars-Sinai Medical Center
Jon Kobashigawa MD
Christine Sumbi

Funding
Irving Precision Medicine Award
NCATS UL1 TR001873