Building the Digital Infrastructure for Clinical Research and AI in Cardiovascular Medicine at Northwestern

Faraz S. Ahmad, MD, MS
Assistant Professor of Medicine (Cardiology)
Assistant Professor of Preventive Medicine (Health and Biomedical Informatics)
Center for Health Information Partnerships
Faraz.Ahmad@northwestern.edu
Twitter: @FarazA_MD
Disclosures

- Received consulting fees from Amgen
- Funded by PCORI, AHRQ, and NHLBI
Outline

1. Bluhm Cardiovascular Institute and the CARdiovascular DIGital AI Core (CARDIAC)

2. Progress and Challenges in Data Plumbing

3. Future Directions in AI in BCVI

4. Resourcing
Bluhm Cardiovascular Institute

Depicts NM sites with cardiovascular presence
US News & World Report Cardiology and Heart Surgery Ranking

Goal: Top 10 by 2020

- Top
- NR = Not ranked

2005 38
2006 NR
2007 19
2008 22
2009 14
2010 16
2011 17
2012 12
2013 13
2014 9
2015 6
2016 7
2017 7
2018 7
2019 1

Bluhm Cardiovascular Institute Launch

Bluhm Cardiovascular Institute today!
Northwestern Medicine Launches New Center Using Artificial Intelligence and Machine Learning to Treat Cardiovascular Disease

Transformative $25 million gift from the Bluhm Family Foundation helps fund new center
Northwestern Medicine EDW Overview

What is the EDW?

• The Northwestern Medicine® Enterprise Data Warehouse (NMEDW) is a joint initiative across the Northwestern University Feinberg School of Medicine (FSM) and Northwestern Memorial HealthCare (NMHC)

• **Mission Statement:**
  • To create a single, comprehensive, and integrated repository of all clinical and research data sources on the campus to facilitate research, clinical quality, healthcare operations, and medical education

• The EDW contains data from 140+ source systems, and is continuously loading data from new source systems

Courtesy of Dan Schneider
Goal is to obtain raw data and labels

- Patient-reported outcomes, mHealth data, sensors
- Cardiac MRI
- Endomyocardial Biopsies
- Pulmonary Function Tests
- Echo
- Curated lists of medications, labs
- ECG
- Cardiac CT
- Cardiac Cath

Demographics, standardized definitions for CVD, co-morbidities, interventions
Use Cases of CARDIAC

**Clinical Research**
- AI/Machine Learning Studies
  - Trial Eligibility
  - Pragmatic Trials
  - Epidemiological Studies

**Patient Care and Operations**
- Population Health Datamarts
- Analytics
- Quality Reporting

**CARDIAC**
- Cardiovascular Digital Artificial intelligence Core

**The SA Node**: 2 NVIDIA TITAN RTX GPUs on HIPAA-compliant on FSM Server
Diverse expertise required to solve data problems

- Clinicians
- Informaticists
- Computer Scientists
- Data Architects and Analysts
- IT Administrators
- Operations, Registry, and Quality Leaders
- Health System Administrators
- Data Privacy and Security Experts
- More...
Outline

1. Bluhm Cardiovascular Institute and the CARdiovascular DIgital AI Core (CARDIAC)

2. Progress and Challenges in Data Plumbing

3. Future Directions in AI in BCVI

4. Resourcing
Different systems before and after 2017

Since 1/2017, 250,000+ echos under 70+ different names and over 1700 unique data elements

Images stored in health system imaging archive
Patient-reported outcomes, mHealth data, sensors

Cardiac MRI

Endomyocardial Biopsies

Pulmonary Function Tests

Exercise and Nuclear Testing

Demographics, standardized definitions for CVD, co-morbidities, interventions

ECG

Cardiac CT

Cardiac Cath

Cardiovascular Digital Artificial Intelligence Core
NMPRO PROGRAM MISSION

Define, expand, and continuously improve the Patient Reported Outcome infrastructure and systems to improve delivery of clinical care and outcomes, engage patients in their care, and accelerate innovation and research.
Heart Failure PROM Pilot Program

• Created process map of clinic workflow with key stakeholders

• Selected and programmed two measures
  – Kansas City Cardiomyopathy Questionnaire – 12
  – PROMIS Global Physical and Mental Health (10 questions)

• Launched pilot in two HF clinics in September 2019
  – Completion via portal/NM app or tablet in clinic

• COVID-19 impacted evaluation, transitioned to NM App/Portal completion
The PROMIS-Plus-HF Profile Measure

**Development**

- In-person focus groups (8 groups; 61 patients) and clinician (n=10) phone interviews to identify content themes
- Mapping of themes to existing PROMIS domains/items and drafting with cognitive testing of new domains/items

**Validation**

- Good psychometric properties (reliability, validity, and responsiveness) in cross-sectional sample of 600 patients and longitudinal sample of 75 patients

**Key Properties**

- Complete assessment of physical, mental, and social health
- Leverages existing, extensively tested PROMIS items with new content specific to heart failure
- Entire instrument or subsets of domains can be used depending on clinical or research purpose

Ahmad, Cella et al. *Circ Heart Fail* 2019.
# The 10-item PROMIS-Plus-HF Clinical Profile

<table>
<thead>
<tr>
<th>Subdomain</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PHYSICAL</strong></td>
<td></td>
</tr>
<tr>
<td>Dyspnea</td>
<td>Over the past 7 days, how short of breath did you get while walking 50 steps/paces on flat ground at a normal speed without stopping</td>
</tr>
<tr>
<td>Fatigue</td>
<td>In the past 7 days, how often were you too tired to do your household chores?</td>
</tr>
<tr>
<td>Physical Function</td>
<td>Does your health now limit you in walking about the house?</td>
</tr>
<tr>
<td>Symptoms</td>
<td>In the past 7 days... Did you feel dizzy or lightheaded?</td>
</tr>
<tr>
<td>Symptoms</td>
<td>In the past 7 days... Did you have swelling in your feet or legs?</td>
</tr>
<tr>
<td>Sleep Disturbance</td>
<td>In the past 7 days... My sleep quality was...</td>
</tr>
<tr>
<td><strong>MENTAL</strong></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>In the past 7 days... I felt that I wanted to give up on everything</td>
</tr>
<tr>
<td>Anxiety</td>
<td>In the past 7 days... I worried that my health would get worse</td>
</tr>
<tr>
<td><strong>SOCIAL</strong></td>
<td></td>
</tr>
<tr>
<td>Ability to Participate in Social Roles and Activities</td>
<td>I am limited in doing my work (include work at home)</td>
</tr>
<tr>
<td>Independence</td>
<td>In the past 7 days... Were you dependent on others to get things done?</td>
</tr>
</tbody>
</table>

Estimated completion time: 2-3 minutes
NLP for Cardiac MRI Reports

Collaborating with McCormick AI students, Hammond, Luo, Naidech Labs, NMEDW

Word embeddings from impressions of 2,363 cMRI reports

80% Training 20% Testing

Normal
Mildly Reduced
Mod Reduced
Sev Reduced

Model accuracy

Note: True accuracy ~90-95% given tested on “noisy” labels

Courtesy of Ramsey Wehbe, MD
Important Question to the Heart Failure Team

• Can we identify patients who are the highest risk in our health system and would benefit from our advanced therapies, such as heart transplant, heart pumps, and investigational devices?
Building Relationships that are Data Driven to Optimize Longevity in Heart Failure

BIRD DOG - HF

• Inclusion
  - At least one code for heart failure
  - LVEF ≤ 40% by echo (most recent)
  - At least one encounter with NM within the 18 months
  - At least on high risk features:
    • ≥ 2 HF hospitalizations, ED visits within one year of last encounter
    • Hospitalization with IV inotropes use within one year of last encounter
    • Most recent Creatinine > 1.6 and < 4.0
    • Most recent Sodium < 133
    • Most recent systolic BP < 90
    • On high dose diuretic as outpatient at any time
    • Absence of beta blocker or RAAS (ACE/ARB) inhibition (currently)
Initial Stats (n=2148)

- Never seen a cardiologist: 738
- Never seen an HF specialist: 1,778
Outline

1. Bluhm Cardiovascular Institute and the CARDiovascular DIgital AI Core (CARDIAC)

2. Progress and Challenges in Data Plumbing

3. Future Directions in AI in BCVI

4. Resourcing
Real-time NLP Pipelines for COVID-19 Case Identification

EDW
Epic Telephone Encounters

Covid Positive
(1,000 notes)

Covid Negative
(4,000 notes)

SA Node

Fine-tune SOTA models

- clinicalBERT
- ULMFiT – forward model
- ULMFiT – backward model

Microsoft Analytics suite

- Data Factory
- Cognitive Services
- Azure SQL DB
- Azure Machine Learning
- Data Catalog
- Cortana
- HDInsight
- Data Lake Store
- Data Lake Analytics
- Bot Framework
Future of AI in home heart failure care at Northwestern

PATIENT

Patient-generated health data (e.g., weight, blood pressure)

Patient-reported outcomes (PROMIS-Plus-HF Clinical Profile)

MyNM App

Electronic Health Record (Epic)

AI-powered algorithms

Clinical Staff/Providers

Courtesy of Anjan Tibrewala, MD, and Jay Pandit, MD
BCVI collaboration with Etemadi Lab
Theoretical Deep Learning Risk Prediction for COVID-19 & Heart Failure

Recurrent Neural Network

Labs

Time Vector $t_0 - t_1$

Modality Vector

Clinical Notes

Time Vector $t_0 - t_3$

Modality Vector

$t=1$

$t=2$

$t=3$

$t=0$

Output: Cox-like survival function

“Stagnant” Features (e.g. Age, demographics)

Imaging (Unstructured images and/or labs)

Courtesy of Ramsey Wehbe, MD
Outline

1. Bluhm Cardiovascular Institute and the CARdiovascular DIgital AI Core (CARDIAC)

2. Progress and Challenges in Data Plumbing

3. Future Directions in AI in BCVI

4. Resourcing
The Dream: The Learning Health System

Aligning with Priorities of Stakeholders with Shared Interests

Bootstrapping CARDIAC over 2 years

• Me (on K12, donated time)
• NMEDW Research Manager: Dan Schneider (donated time)
• NMEDW Data Architect (10% on sponsored project)
• NMEDW Research Analyst (20% on sponsored project/my start-up)
• NM BCVI Architect and Analyst (contribute time around specific projects)
• NM Information Services
  – Research and Development Team
• Yuan Luo, Andrew Naidech (donated time) and their pre-doctoral students
• Several additional BCVI collaborators: Sanjiv Shah, Kannan Mutharasan, Shruti Cruz
• Med students, fellows
• PROs
  – K12 Team, NM PRO Steering Committee, and the NM Quality Innovation Center
People and Data are key to the CV Informatics Program

- A curated, updating CV data core will be the foundation for the Center for AI and CVD and many larger BCVI and NM initiatives.
- Diverse expertise and creativity are needed:
  - Clinicians
  - Informaticists
  - Computer Scientists
  - Data Architects and Analysts
  - IT Administrators
  - Operations, Registry, and Quality Leaders
  - Health System Administrators
  - Data Privacy and Security Experts
  - More...

**The SA Node:** A HIPAA-compliant, 2 GPU, 15 CPU Core on FSM Server
Proposed Staffing for Cardiovascular Informatics Program

- Leadership: Clinical, research, and informatics expertise
- Program Manager, ideally with informatics background
- 1-2 Data Engineers embedded with NMEDW and NM R&D teams
- 1 Data Analyst
- Funding for informatics pre- and post-docs
- Center for AI and CVD AI fellows
- Continued partial support (financial or informatics) of collaborators
Acknowledgements

- Abel Kho
- Patrick McCarthy
- Sanjiv Shah
- Kannan Mutharasan
- Jay Pandit
- Anjan Tibrewala
- Shruti Cruz, MBA
- Ramsey Wehbe
- Dan Schneider
- Martin Borsje
- Aditya Sharma
- Anna Pawlowski
- Kris Hammond
- Yuan Luo
- Mozzi Etemadi
- NM Research and Development Team
- Han Liu, Kevin Yu, Ted Hu
- Andrew Naidech
- Lacey Gleason
- David Cella
- Karl Bilimoria
- Nan Rothrock
- Kristina Davis
- Clyde Yancy
- Susan Yount
- Martha-Margaret Cotton
- Leilani Lacson
Questions?

Faraz.Ahmad@northwestern.edu
Twitter: @FarazA_MD