

Strategic Health IT Advanced Research Projects (SHARP) and Beacon Communities for HIT standards and interoperable data exchange

Christopher G. Chute, MD DrPH Mayo Clinic Successfully Leveraging HITECH Resources MN eHealth Summit 17 June 2010



### Biomedical Informatics

# SHARP: Area 4: Secondary Use of EHR Data A \$15M National Consortium

- 14 academic and industry partners
- Develop tools and resources that influence and extend secondary uses of clinical data
- Cross-integrated suite of project and products
  - Clinical Data Normalization
  - Natural Language Processing (NLP)
  - Phenotyping (cohorts and eligibility)
  - Common pipeline tooling (UIMA) and scaling
  - Data Quality (metrics, missing value management)
  - Evaluation Framework (population networks)

© 2009 Mayo Clini

#### Biomedical Informatics

# Collaborations

- Agilex Technologies
- CDISC (Clinical Data Interchange Standards Consortium)
- Centerphase Solutions
- Deloitte
- Group Health, Seattle
- IBM Watson Research Labs
- University of Utah

- Harvard Univ. & i2b2
- Intermountain Healthcare
- Mayo Clinic
- Minnesota HIE (MNHIE)
- MIT and i2b2
- SUNY and i2b2
- University of Pittsburgh
- University of Colorado

### Biomedical Informatics

# **Project Advisory Committee**

Suzanne Bakken, RN DNSc, Columbia University
C. David Hardison, PhD, VP SAIC
Barbara A. Koenig, PhD, Bioethics, Mayo Clinic
Issac Kohane, MD PhD, i2b2 Director, Harvard
Marty LaVenture, PhD MPH, Minnesota Department of Health
Dan Masys, MD, Chair, Biomedical Informatics, Vanderbilt University
Mark A. Musen, MD PhD, Division Head BMIR, Stanford University
Robert A. Rizza, MD, Executive Dean for Research, Mayo Clinic
Nina Schwenk, MD, Vice Chair Board of Governors, Mayo Clinic
Kent A. Spackman, MD PhD, Chief Terminologist, IHTSDO
Tevfik Bedirhan Üstün, MD, Coordinator Classifications, WHO

#### Biomedical Informatics

# Vision

- To assemble a federated informatics research community committed to open-source resources that can industrially scale to address barriers to the broad-based, facile, and ethical use of EHR data for secondary purposes
- To create, evaluate, and refine informatics artifacts that advance the capacity to efficiently leverage EHR data to improve care, generate new knowledge, and address population needs
- To make these artifacts available to the community of secondary EHR data users, manifest as open-source tools, services, and scalable software
- To partner with industry developers who can make these resources available with commercial deployment and support.

Themes & Projects						
Themes		s	Projects	Players		
ation	Phenotype Recognition	Data Quality and Evaluation Frameworks	Clinical Data Normalization	IBM, Mayo, Utah, Agilex		
Data Normaliza			Natural Language Processing (NLP)	Harvard, Group Health, IBM, Utah ,Mayo, MIT, SUNY, i2b2, Pittsburgh, Colorado		
			High-Throughput Phenotyping	CDISC, Centerphase, Mayo, Utah		
			UIMA and Scaling Capacity	IBM, Mayo		
			Data Quality	Mayo, Utah		
			Evaluation Framework	Agilex, MN HIE, Mayo, Utah		

### Biomedical Informatics

### Project 1 - Clinical Data Normalization CG Chute, MD DrPH

- Build generalizable data normalization pipeline
- · Semantic normalization annotators involving LexEVS
- Establish a globally available resource for health terminologies and value sets
- Establish and expand modular library of normalization algorithms

# 

Project 2: Clinical Natural Language Processing (cNLP); GK Savova, PhD

- Overarching goal
  - High-throughput phenotype extraction from clinical free text based on standards and the principle of interoperability
- Focus
  - Information extraction (IE): transformation of unstructured text into structured representations
  - Merging clinical data extracted from free text with structured data

# Biomedical Informatics

# Project 3: High-Throughput Phenotyping Jyoti Pathak, PhD

- Develop portable phenotype algorithms
  - Administrative Data; Labs; Medications
  - Build on NHGRI eMERGE experience
- Phenotyping logic specification
- Applications of phenotype characterization
- Expansion of Cohort Amplification methods

#### Biomedical Informatics

### Project 4 - UIMA exploitation Marshall Schor – IBM Research

- Use UIMA as a unifying framework, leveraging ecosystem
  - Work with team leads to identify "fit" (or not) of UIMA into subprojects
    - Phenotyping and Data Quality, especially
- Support UIMA and UIMA-AS use
  - Consult on pipe line design / architectures / configuration
- Support scaling, capacity flexibility
  - Develop and deploy virtual machine images that can dynamically scale in cloud computing environments
  - Develop integration / deployment tooling with goal of simplicity
    - Enabling widespread adoption of POC

### Biomedical Informatics

04/29/10

# Project 5 - Data Quality Kent Bailey, PhD

© 2010 Mayo Clinic

- Refine metrics for data consistency
- Deploy methods for missing or conflicting data resolution
- Integrate methods into UIMA pipelines
- Refine and enhance methods

# Biomedical Informatics

Project 6 - Real-world evaluation framework Dr. Huff

- Iteratively test normalization pipelines, including NLP where appropriate, against normalized forms, and tabulate discordance.
  - Normalize retrospective data from the EMRs and compare it to normalized data that already exists in our data warehouses (Mayo Enterprise Data Trust, Intermountain).
- Use cohort identification algorithms in both EMR data and EDW data.
  - Normalize the data against CEMs.



### Biomedical Informatics



• Disease metrics, utilization of ER and hospital

© 2009 Mayo Clinic



#### Biomedical Informatics Biomedical Informatics Beacon Grant for SE MN Goals **Childhood Asthma and Diabetes** Vision Reduce Emergency room visits SE MN community will embrace standards based HIE to improve access, quality and efficiency of Reduce unscheduled MD visits health care delivery Reduce hospitalization •MN HIF Improve self-reported functioning •County Public Health •Stratis – REACH (HIT Ext.) Departments Improve compliance with the treatment •Key Health Alliance of asthma •Mayo Clinic Rochester Agilex Mayo Health System Improve school attendance Veterans Administration •Olmsted Medical Center • Reduce days out of work – self-reported for Diabetes Indian Health Service • Public school systems Improve compliance with Diabetes • US DOJ – Bur. Prisons •Winona Health System

Biomedical Informatics	SE MN Beacon	
EMRs	Care Provision	<u>Counties</u>
OMC	Public health	Dodge county
MCR	Nursing homes	Fillmore county
MHS	Hospitals	Freeborn county
WHS	Emergency rooms	Goodhue county
	Home health	Houston county
	Schools	Mower county
	Out-patient clinics	Olmsted county
		Rice county
		Steele county
		Wabasha county
		Winona county

### Biomedical Informatics

# Population Health Improvement Goals and Needs for Beacon

- Identify affected individuals currently without effective treatment
- Engage public health services for underserved populations
- Improve telemedicine connections in rural communities through Winona's eSuite
- Report score card every month to practices on performance
- Provide platform for analysis of data for comparative effectiveness research





