MAVERIC \informatics

MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

Leonard D'Avolio, PhD
Assoc. Center Director,
Biomedical Informatics
MAVERIC, VA Boston
Harvard Medical School

Introduction

- MAVERIC
  - 140+ **multi-disciplinary** research & development
    - Epidemiology (13 yrs)
    - Biospecimen repository (11 yrs)
    - Large scale clinical trials (8 yrs)
      - ISO Certified
    - Informatics (3.5 yrs)
- Our vision is to create a *Learning Healthcare System* within VA through application of research resources and methodologies to important clinical problems.

MAVERIC\informatics

Transform healthcare through innovative uses of information

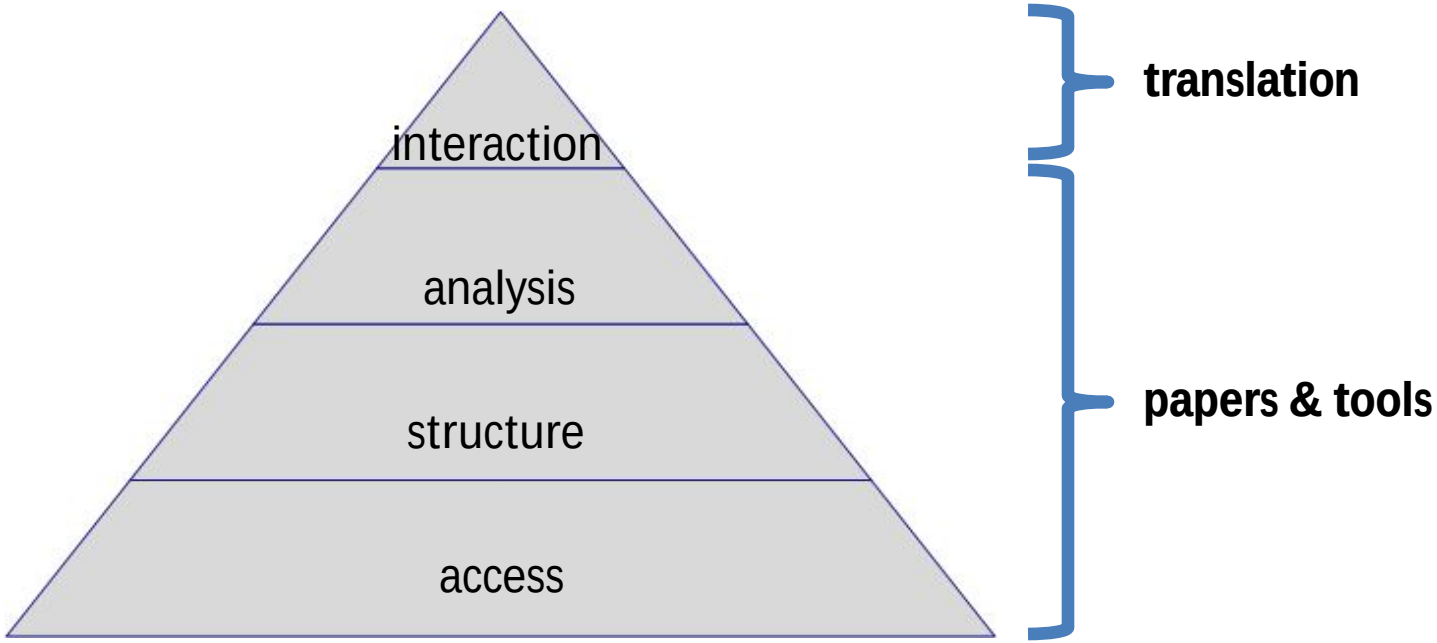
- ~40 team members

-11 developers
-3 database
-5 systems analysts
-2 QA
-3 PMs

-2 system admins
-Dir of Bioinformatics
-Dir of Operations
-10 contractors
-3 post-docs

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

Approach



D'Avolio LW, Farwell WR, Fiore LD. 2010 Dec;123(12 Suppl 1):e32-7. **Comparative effectiveness research and medical informatics.** Am J Med.

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

Approach

- Problems from clinical collaborators
- National in scope
- No “one-off” solutions
- Research methodology + development best practices

Challenge:
Facilitate ‘omic discovery

- ‘omic science requires enormous sample sizes
  - Disease more granular
  - Free rider dilemma
  - Disconnected care / research systems
- Results
  - Underpowered studies
  - Limits discovery

The VA's Unique Opportunity

- Learn how to keep Veterans healthy
- Maximize investment in genomic science
  - Genomic discovery requires large sample sizes
- Intramural research program
- Advance medical knowledge for society
  - 6 million "active" users
  - 20+ years of EMR data
- *83% of Veterans support genomic database
  - 71% would definitely or probably participate

*Kaufman et al. Veterans attitudes regarding a database for genomic research. Genetics in Med. (2009) 11, 329-337

The Million Veteran Program

The Million Veteran Program (MVP)

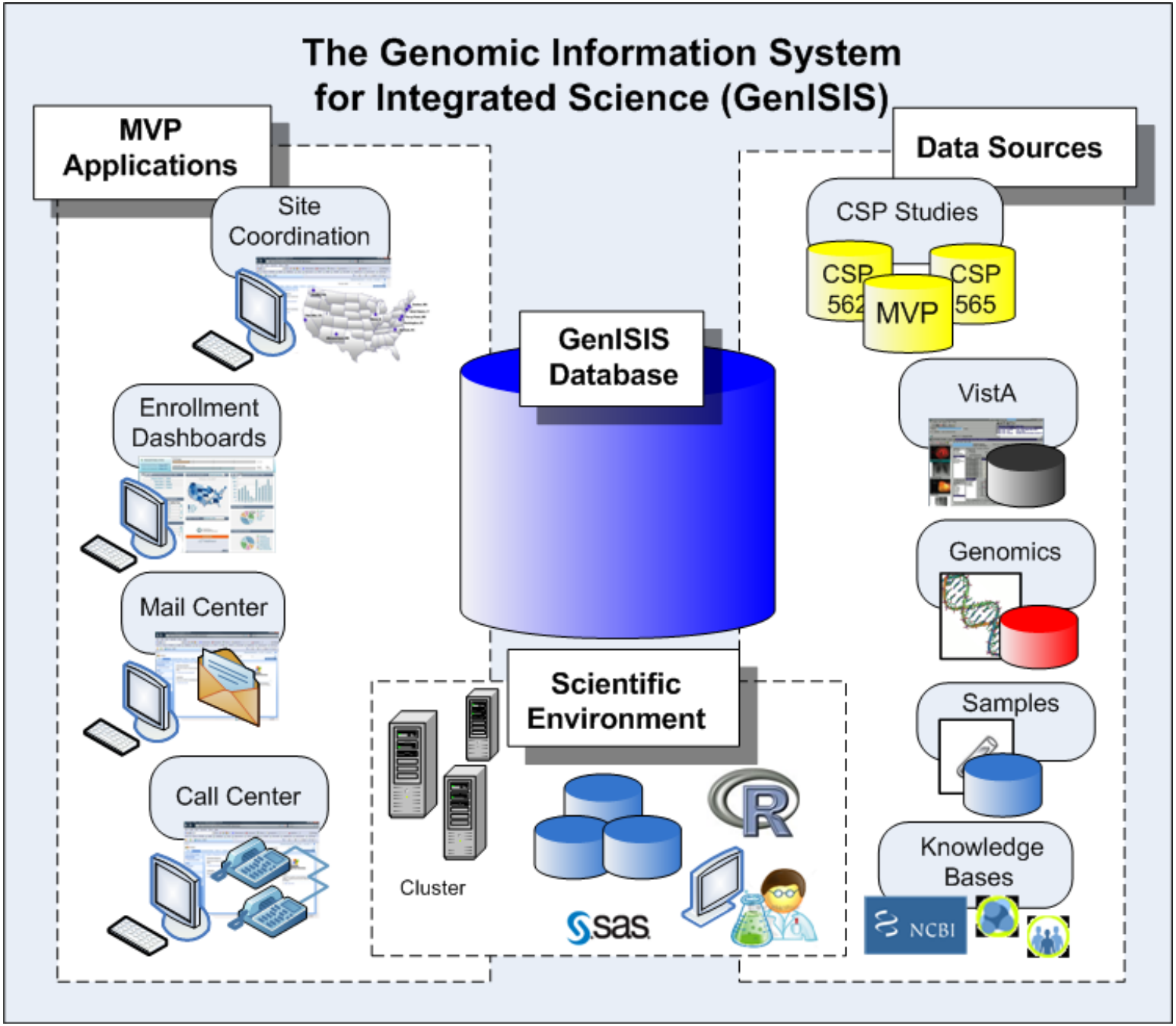
- The goal: 1m Veteran volunteers in 5-7 years
- Survey
  - 5 page baseline
  - 15 page comprehensive
- Blood sample
- Open consent & HIPAA authorization
- Access to the medical record
- Ability to re-contact

MVP Logistics

- 40 facilities enrolling
  - Scaling to 55
- Mail to every Veteran at enrolling clinics
- Survey & scheduling preferences by mail
- Consent & blood draw in person
- Call center for questions

Heavily Automated

- 20k invitations a week
- 7 different mail types
- 2000 calls per week (inbound + outbound)
- Dynamic form generation at sites
- 353 unique reports for 235 users
- Unified view of all interactions with Veteran


The Genomic Information System
for Integrated Science (GenISIS)
MVP
Applications
Site
Coordination
Enrollment
Dashboards
Mail Center
Call Center
GenISIS
Database
Scientific
Environment
Cluster
sas
Data Sources
CSP Studies
CSP
562
MVP
CSP
565
VistA
Genomics
Samples
Knowledge
Bases
NCBI

Secure Scientific Environment

- High performance computing environment
- Accessible to VA-credentialed investigators
- Data & analysis within the VA firewall
- Working on governance / access policies
  - *VA CSP DNA Bank founded May 1999

*Lavori et al. Principles, organization, and operation of a DNA bank for clinical trials: a Department of Veterans Affairs cooperative study. Con Clin Trials (2002) 23, 222-239

Veterans Health Administration
Research & Development
Improving Veterans' Lives -> www.research.va.gov

GenISIS C3
(Welcome VHA01\vhabhssrinas) Feedback

Patient Selector
GenISIS ID: Search Clear
Patient: last_name02, first_name02

Home
Patient
Site
Visit
Mailing
Admin
Dashboards

Enrol | Contact | Validation | PDF Form Test

Informed Consent Form
Download Form ☑ Completed/Signed
HIPAA Authorization
Baseline Survey Form
Blood Collection

© 2010 GenISISC3 Link 1 Link 2 Link 3

http://vhabosdev7/publishGenISISC3/Visit/DownloadForm.aspx?file=../bin/Visit/MVP_Informed_Conse - Microsof...

File Edit Go To Favorites Help

Page Tools

1 / 9 78.8% Find

VA Department of Veterans Affairs	VA RESEARCH CONSENT FORM Page 1 of 8

Participant Name: first_name02 last_name02 10000002 Date: 5/12/2010

Title of Study: Million Veteran Program (MVP) (CSPG002)
Principal Co-Investigators: Facility:

INTRODUCTION
You are being invited to take part in a research project that is being funded by the Department of Veterans Affairs. Before you decide to take part, it is important for you to know why the research project is being done and what it will involve, to include any potential risks to you as well as any potential benefits you may receive.

Read the information below closely, and discuss it with family and friends if you wish. Ask one of the project staff if there is anything that is not clear, or if you would like more details. Take your time to decide. If you do decide to take part, your signature on this consent form, will show that you received all of the information below and that you were able to discuss any questions and concerns you had with a member of the project team.

612.0 x 792.0 pt

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

GenISIS Command and Control Center - Microsoft Internet Explorer provided by VA New England Health Care

http://localhost:2961/Visit/PDFFormtest.aspx

File Edit View Favorites Tools Help

GenISIS Command and Control Center

Veterans Health Administration
Research & Development
Improving Veterans' Lives www.research.va.gov

GenISIS C3
(Welcome VHA01\vhabhssrinas) Feedback

Patient Selector
GenISIS ID: Search Clear
No patient selected

- Home
- Patient
- Site
- Visit
- Mailing
- Admin
- Dashboards

CheckIn | Contact | Enroll | PDF Form Test

- Informed Consent Form
- HIPAA Authorization
- Baseline Survey Form
- Blood Collection

Check List

Time Since Last Meal:
- Unsure
- Less than 8 hours
- 12-16 hours
- Greater than 16 hours

Barcode scan of empty blood tube: 0010006251672

Person obtaining blood:
- Research Coordinator
- Research Assistant
- Research Phlebotomist
- Phlebotomy Lab
- Other

Barcode scan of filled blood tube: 0010006256875

Barcodes do not match,please select a reason: Tube malfunction
- Tube malfunction
- Blood had to be re-drawn
- Other

Save Submit

Done Local intranet 100%

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

GenISIS Investigator - Microsoft Internet Explorer provided by VA New England Health Care

UNITED STATES
DEPARTMENT OF VETERANS AFFAIRS

GenISIS

Investigator | Home | Studies | Data | Analysis Tools | Knowledge Base | eVA | Account | Exit

GenISIS Query & Analysis — Project: i2b2 Demo — User: testPrincipal — Find Patients | Analysis Tools | Help

Navigate Terms | Find Terms

- Male
- Income
- Race
- Disease History
  - Cancer
  - Circulatory
    - Congestive heart failure
    - Coronary artery/Coronary heart disease (includes angin
    - heart attack
    - High blood pressure (Hypertension)

Workplace

- SHARED
- testPrincipal

Previous Queries

- 35-44-Coron-Male@11:10:27 [4-24-2012] [testPrincipal]
- 18-34-Coron-Male@11:09:35 [4-24-2012] [testPrincipal]
- my cohort [4-24-2012] [testPrincipal]
- 35-44 y-High bl@10:39:40 [4-24-2012] [testPrincipal]
- Cancer@09:52:19 [4-24-2012] [testPrincipal]

Query Tool

Query Name: 35-44-Coron-Male@11:10:27

Temporal Constraint: Treat all groups independently

Group 1	Group 2	Group 3
Dates | Occurs > 0x | Exclude	Dates | Occurs > 0x | Exclude	Dates | Occurs > 0x | Exclude
Treat independently	Treat independently	Treat independently
35-44 years old	Coronary artery/Coronary h	Male
45-64 years old		
55-64 years old		
one or more of these	one or more of these	one or more of these

AND AND

Run Query | Clear | Print Query | 3 Groups | New Group

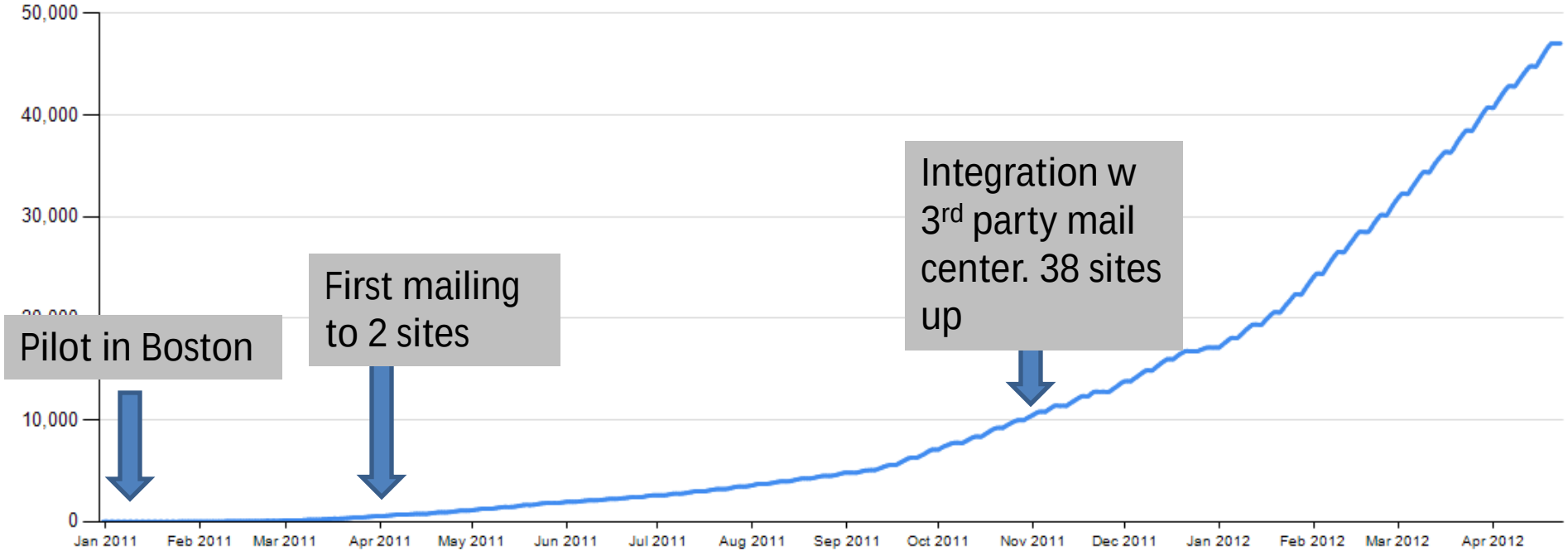
Query Status

Finished Query: "35-44-Coron-Male@11:10:27" [4.1 secs]
Compute Time: 2.1 secs

Number of patients for "35-44-Coron-Male@11:10:27"
patient_count 3333

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

MVP Enrollment to Date



MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

Challenge:
Create New Models of Science

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

STANFORD
MEDICINE

Current models of science do not support clinical effectiveness research

- RCTs too expensive
  - Millions of $ & several years to answer few questions
  - Questions of generalizability
- Observational studies suffer from bias
  - Confounding by indication
  - Data quality issues

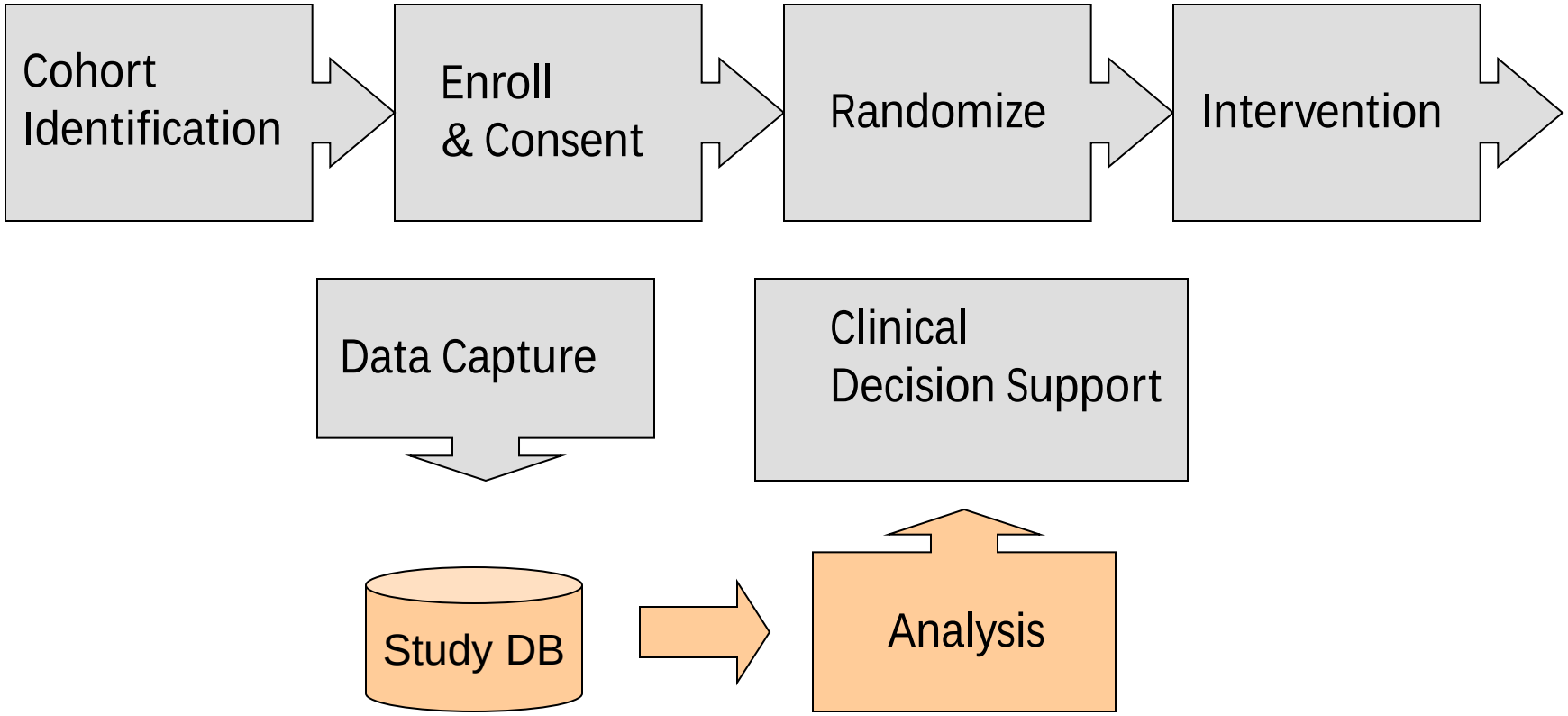
Solution:

- Embed research methods into routine care
  - Turn every equipoise situation into a learning opportunity
- Result
  - Locally “selfish” research & evidence
  - Convert results into decision support
  - Speed translation

Point of Care Research - Clinical Trial Example

- A clinical trial with a substantial portion of its operations conducted by clinical staff in the course of providing patient/subject's routine clinical care and where the choice of treatment is between two "equivalent" options

Care providers using EMR



Study team using traditional scientific tools

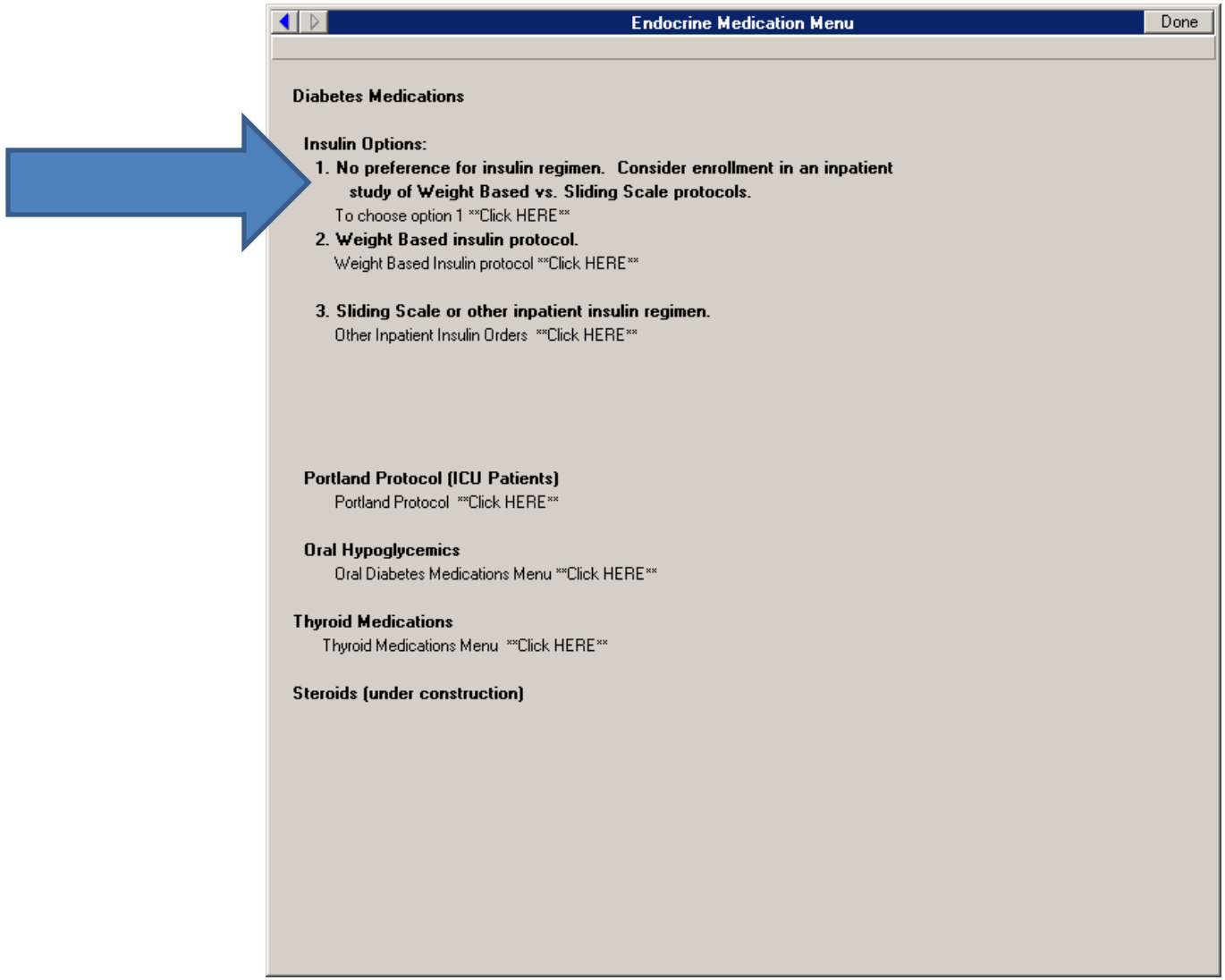
MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

POCR Pilot Study goals

- Establish feasibility of POCR
  - Ability to modify the EMR
  - Physician and patient acceptance
  - IRB and regulatory acceptance
- Settle a substantive clinical issue
- Demonstrate closing the implementation gap

POCR Use Case

- Boston VA initiative
- Insulin protocol
  - Sliding scale insulin regimen
  - Weight based insulin regimen
- Both regimens are approved and in use at VA Boston
- No published data comparing outcomes


Endocrine Medication Menu
Done
Diabetes Medications
Insulin Options:
1. No preference for insulin regimen. Consider enrollment in an inpatient study of Weight Based vs. Sliding Scale protocols.
To choose option 1 **Click HERE**
2. Weight Based insulin protocol.
Weight Based Insulin protocol **Click HERE**
3. Sliding Scale or other inpatient insulin regimen.
Other Inpatient Insulin Orders **Click HERE**
Portland Protocol (ICU Patients)
Portland Protocol **Click HERE**
Oral Hypoglycemics
Oral Diabetes Medications Menu **Click HERE**
Thyroid Medications
Thyroid Medications Menu **Click HERE**
Steroids (under construction)

Recruitment Summary

Recruitment	N (%)
Number of Eligible Patients	129
No response from clinician	17
Clinician refusal	21
Patients who declined participation	4
Patients not enrolled for administrative reasons	4
Number of Patients Enrolled	83 (64.3%)
Patients Randomized	75
Patients Consented to Chart Review	8
Clinician Participation Clinician-Initiated Consults	37 (28.7%)

POCR Requirements

- EMR requirements
  - Incorporate study logic and workflows
  - Feature I/O capability
  - Decision support
- Culture change for patients and providers
- Change in regulatory requirements
  - Informed consent
  - Engagement in research
  - Adverse event reporting

Promising for Biomarker Validation

- Large numbers of candidates that iterate frequently
- Require validation in large populations
- FDA and CMS regulatory and reimbursement questions
- No business model in place for commercialization
  - Why invest millions for an RCT for one-time $100 biomarker?
- Market vacuum now occupied by web businesses

Next Challenge:
Medicine's information dilemma

Medicine's Information Dilemma

Supply

Data

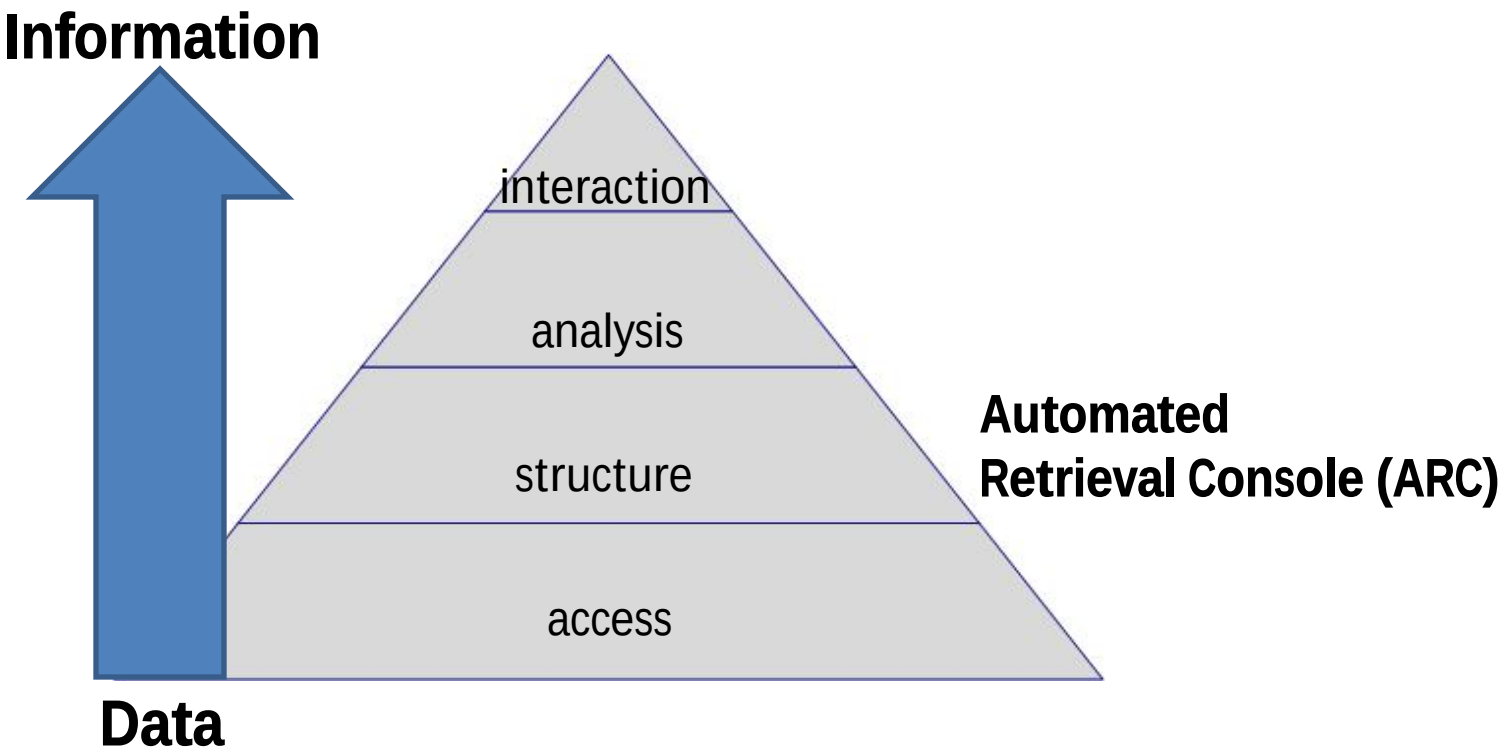
?

Demand

Information

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY RESEARCH AND INFORMATION CENTER

Medicine's Information Dilemma



MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

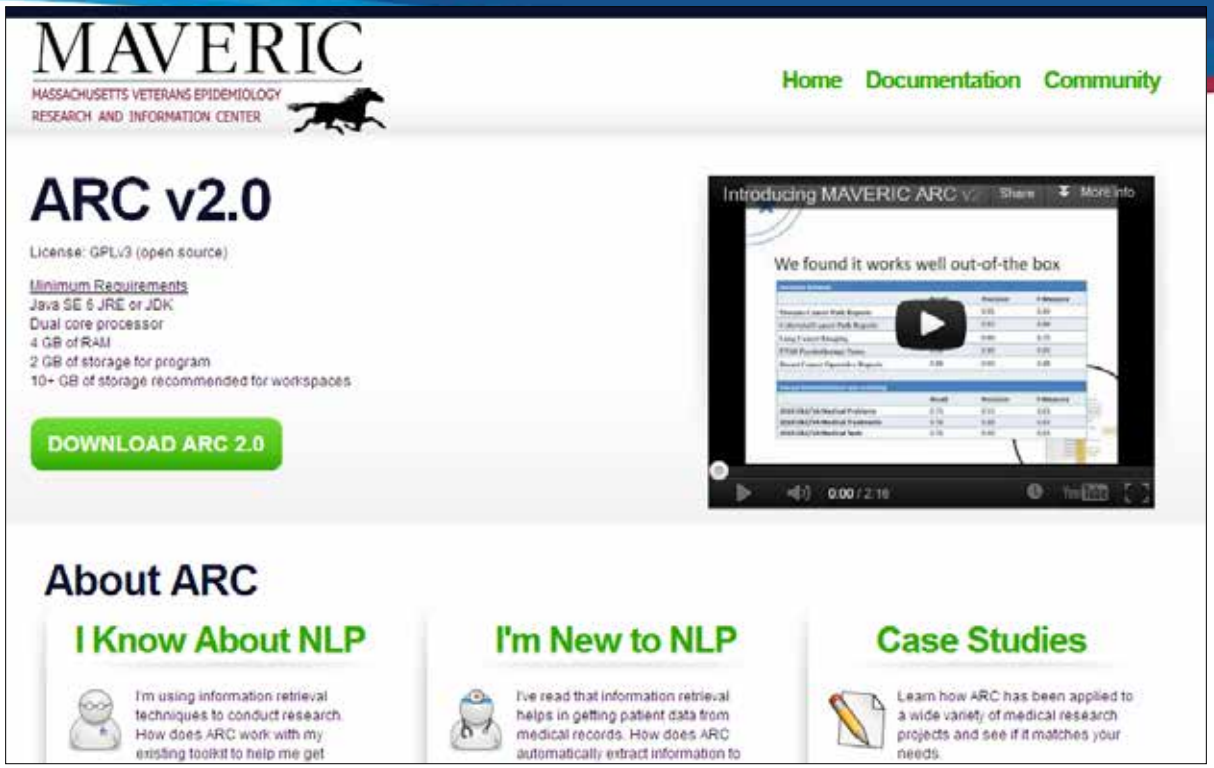
Current approaches to text processing

- Largely rules-based
- Custom solutions
- Require post-processing logic
- Require NLP expertise
- Bottom line - costly & doesn’t scale

The Automated Retrieval Console (ARC)

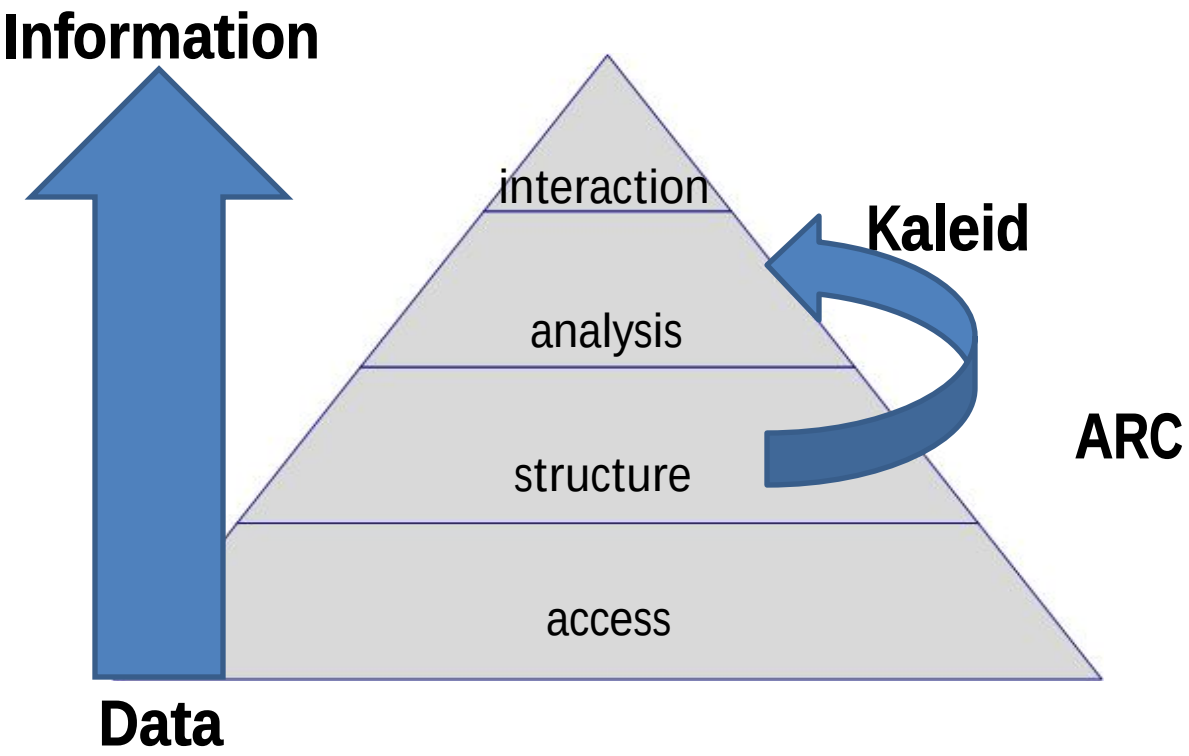
- Find cases or concepts “like this”
- Learns by example
- Reduce entire process to 4 button clicks
- No code / no developers
- 90/90 goal

Document Retrieval			
	Recall	Precision	F-Measure
Prostate Cancer Path Reports	0.97	0.95	0.94
Colorectal Cancer Path Reports	0.90	0.92	0.89
Lung Cancer Imaging	0.76	0.80	0.75
PTSD Psychotherapy Notes	0.98	0.90	0.93
Breast Cancer Operative Reports	0.88	0.90	0.88
Concept Retrieval (inexact span matching)			
	Recall	Precision	F-Measure
2010 i2b2/VA Medical Problems	0.75	0.93	0.83
2010 i2b2/VA Medical Treatments	0.76	0.89	0.82
2010 i2b2/VA Medical Tests	0.76	0.90	0.83



- 150+ downloads (some CTSA sites)
- Hosted on UCSD's iDASH : http://idash.ucsd.edu
- Hosted on VA's VINCI
- Download from : http://arc.4thparadigm.org/

Medicine's Information Dilemma



MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

Current Solution – Business Intelligence

- Business Data
  - Highly structured & fairly homogeneous
  - Fiscal year & quarter
  - Off line analytics
- Data warehousing & reporting
  - Time relative to calendar
  - No provenance
  - Complex queries can be costly
  - Limited content & context
    - Available structured data

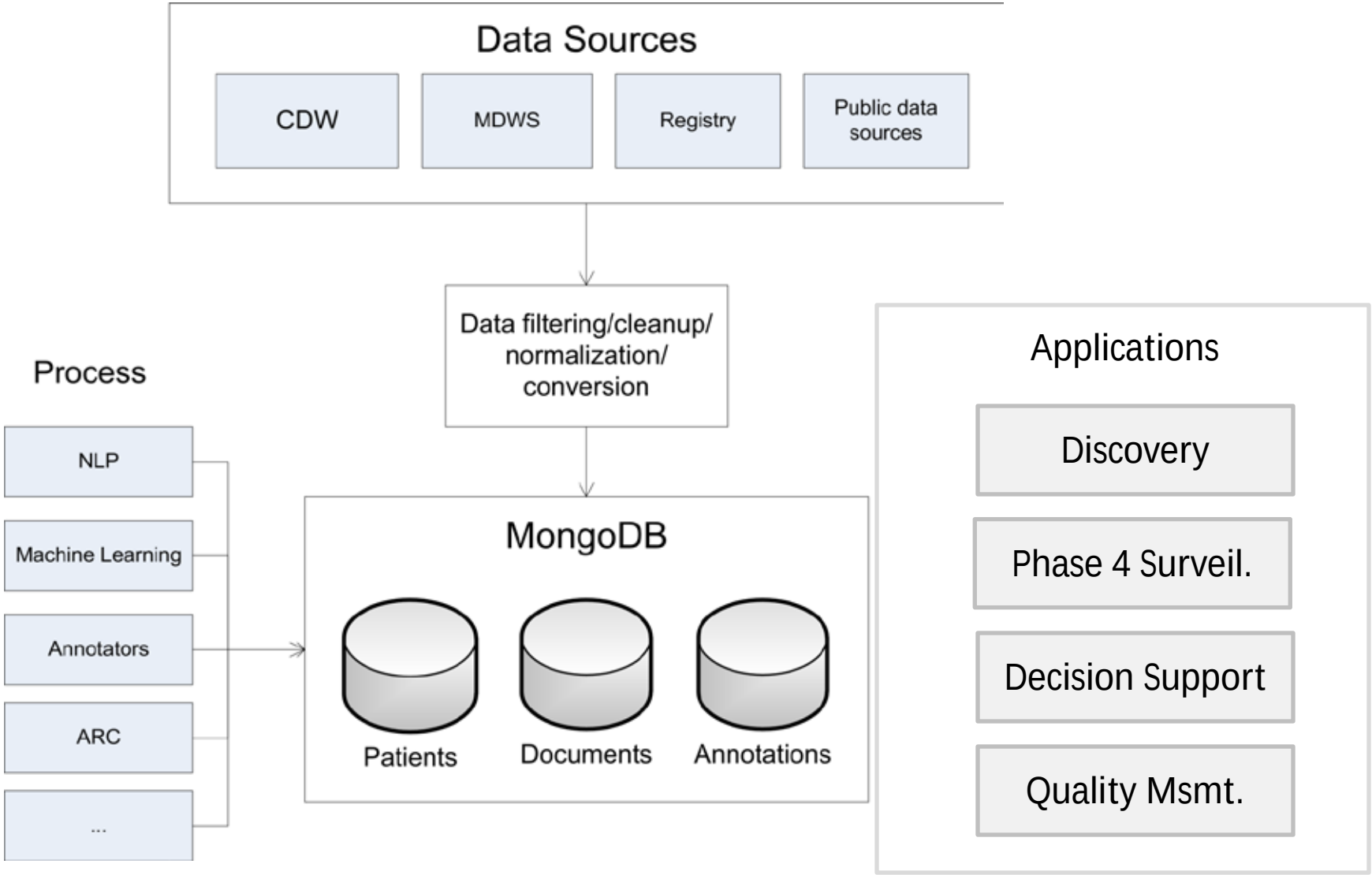
Healthcare Intelligence Requirements

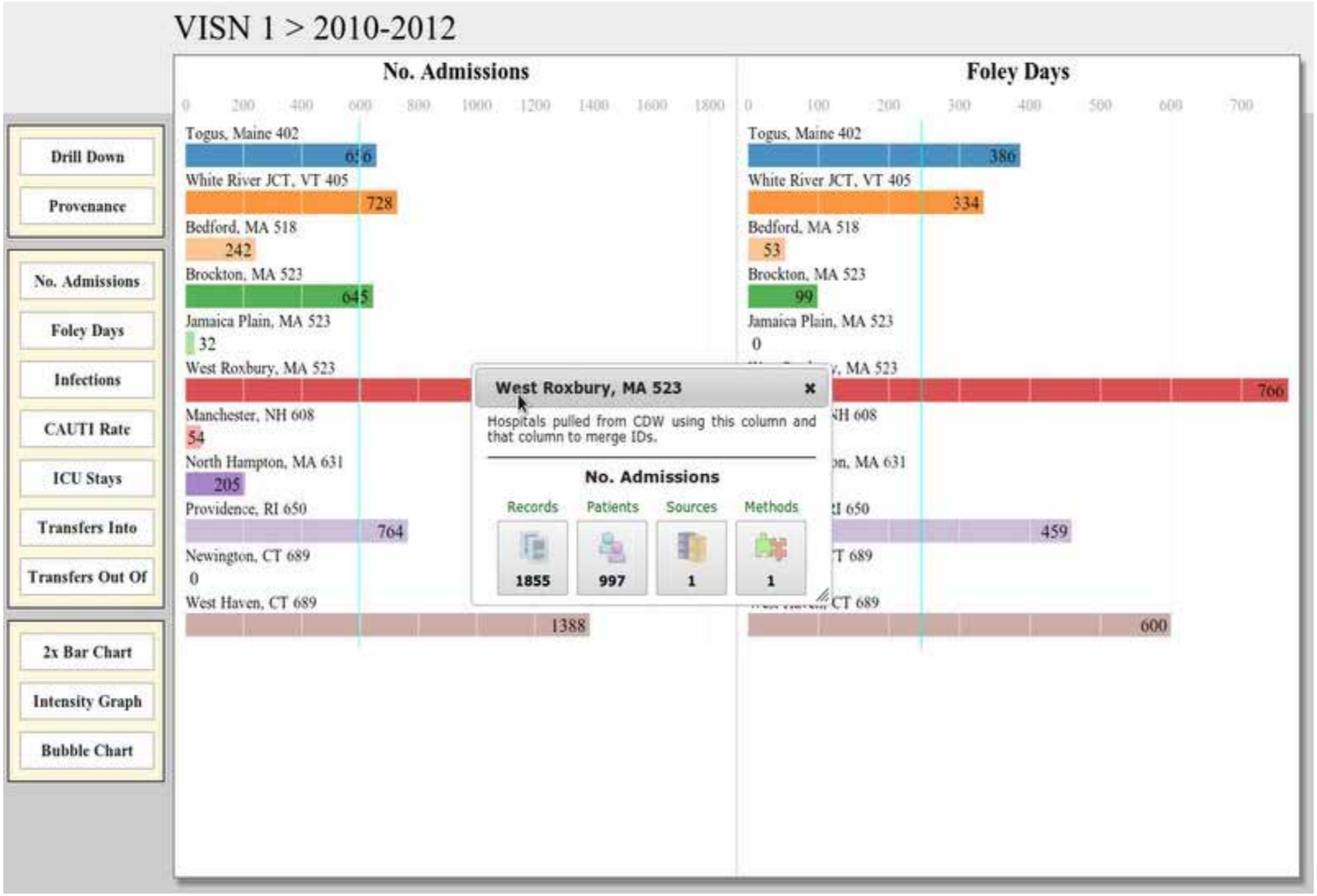
- All data types (including biology)
- Updated regularly from all available sources
- Processes that generalize, not content
- Quantifiable
- Time relative to events
- Near real-time
  - Pre-computed models, pre-processing heavy
- Accessible through & to services
- Robust & accessible provenance

Kaleid Project

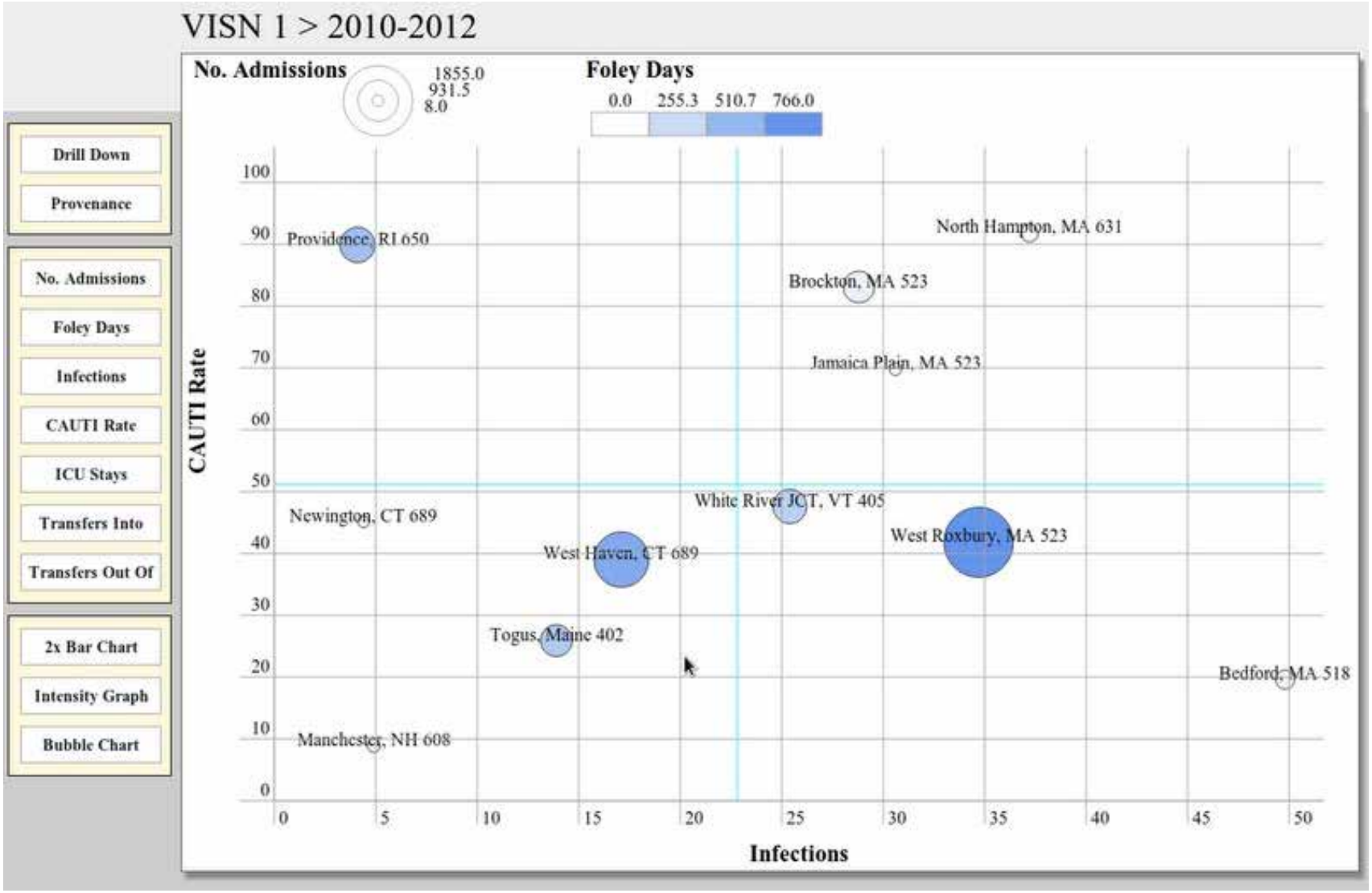
- Domain specific content (not processes)
- Data from all available sources
- Dynamic exploration, real time answers
- Careful provenance
- Domain specific content, generalizable processes

Kaleid Project



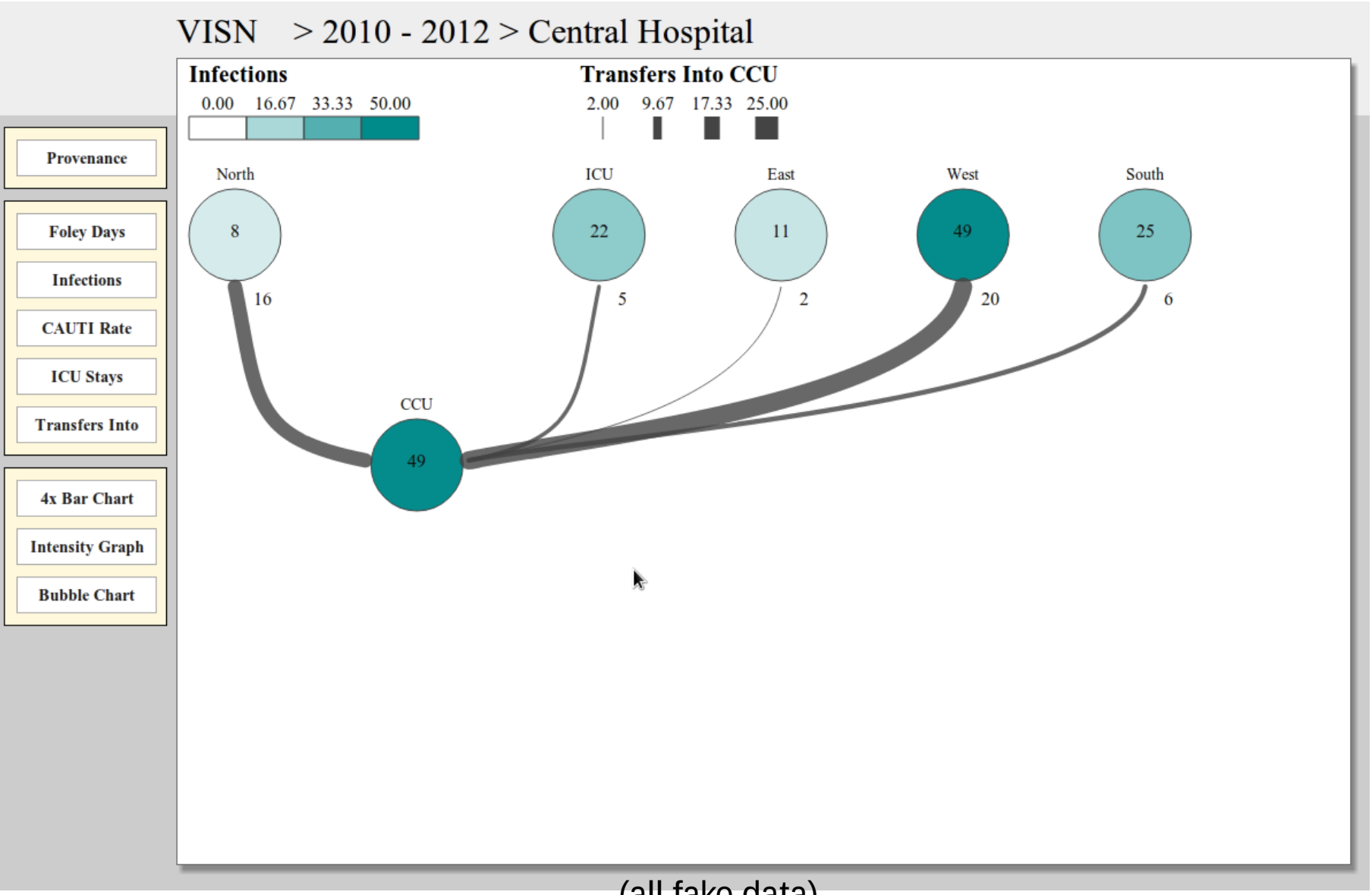


(all fake data)

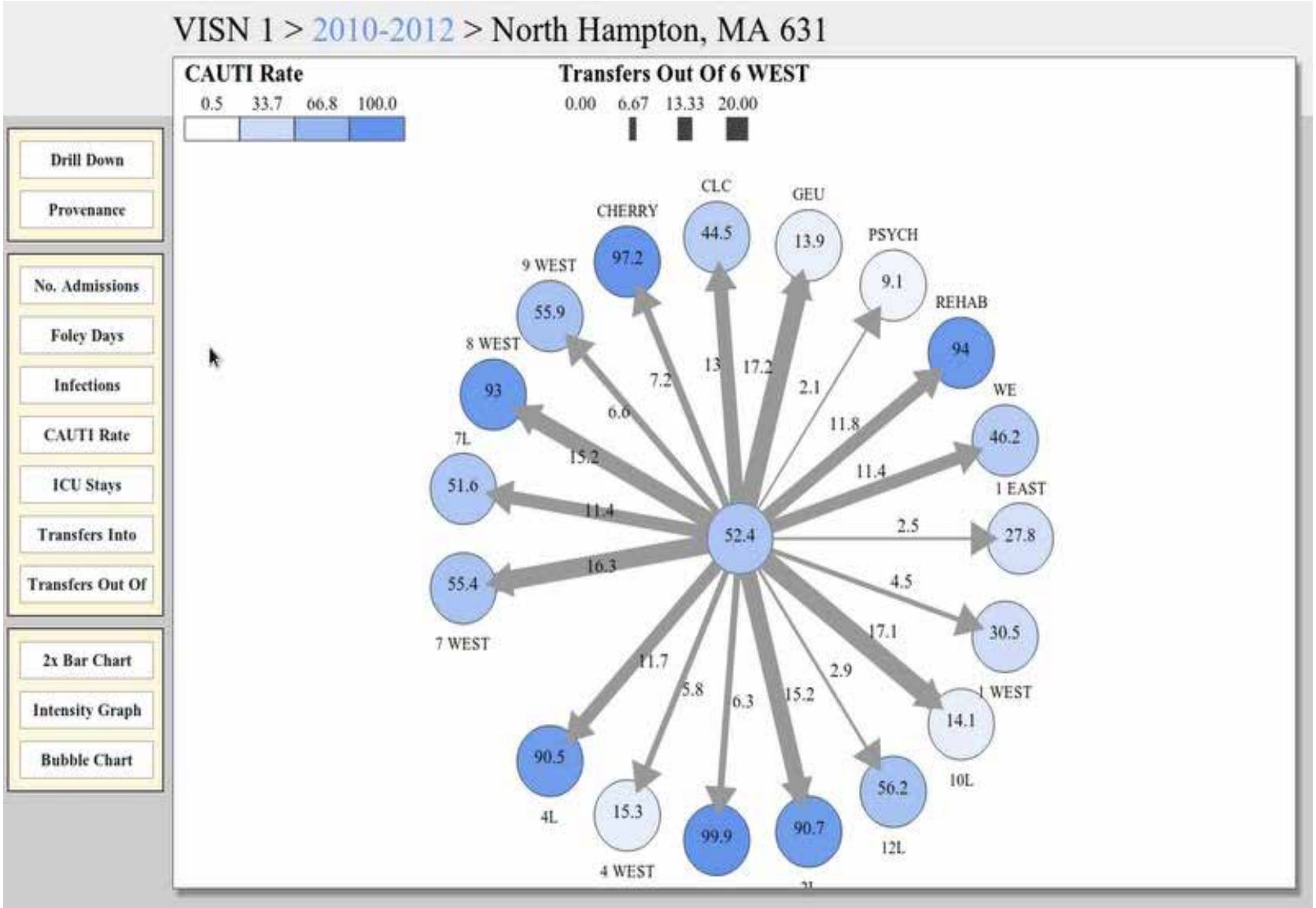


(all fake data)

VISN > 2010 - 2012 > Central Hospital



(all fake data)



(all fake data)

A small subset of the MVP Team

- MVP co-PIs: Mike Gaziano & John Concato
- MVP Director: Colleen Shannon
- Informatics: Len D'Avolio, John Gargas, Saiju Pyarajan
- MAVERIC Exec Director: Louis Fiore
- VA Office of Research and Development
- And many others at MAVERIC & the VA West Haven Clinical Epidemiology Research Center (CERC)

VA POCR Team

- Principal Investigators: Louis Fiore and Philip Lavori
- Co-Investigators: Mary Brophy, James Kaufman, Mike Gaziano
- Informatics: Leonard D'Avolio and Chester Conrad
- CPRS Engineers: Gus O'Neil and Tom Sabin
- Ethics and Informed Consent: John Hermos
- Content Expert: Stephen Swartz
- Data Management: Ryan Ferguson, Galena Sokolovskaya
- Statisticians: Robert Lew, Gheorghe Doros

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

Next-Gen Phenotyping Team

- Leonard D'Avolio
- Frank Meng
- Valmeek "Vick" Kudesia
- Thien Nguyen
- Sergey Goryachev
- Victor Tam

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

Acknowledgements

MAVERIC is sponsored by the

VA Cooperative Studies Program

VA Health Services Research & Development
Office of Research & Development
Department of Veteran's Affairs

For POC-CT: additional NIH support to Stanford University is gratefully acknowledged

CTSA UL1 RR025744

CCSG P30 CA124435

MAVERIC
MASSACHUSETTS VETERANS EPIDEMIOLOGY
RESEARCH AND INFORMATION CENTER

more information

- **Informatics in General**
  - **Comparative Effectiveness Research and Medical Informatics**
    AJM 2009, Volume 123, Issue 12, Pages e32-e37. LW D'Avolio, WR Farwell, LD Fiore
  - **Electronic Medical Records at a Crossroads: Impetus for Change or Missed Opportunity?**
    JAMA 2009;302(10):1109-1111. doi:10.1001/jama.2009.1319. LW D'Avolio

- **Point of Care Clinical Trials**
  - **A Point-of-Care Clinical Trial Comparing Insulin Administered Using a Sliding Scale Versus a Weight-Based Regimen**
    Clin Trials 2011 vol. 8 no. 2 183-195. LD Fiore et al.
  - **Implementation of the Department of Veterans Affairs' First Point-of-Care Clinical Trial**
    J Am Med Inform Assoc 2012 doi:10.1136/amiajnl-2011-000623 (online first). LW D'Avolio et al

- **Future of 'OMIC medicine**
  - **Detours on the Road to Personalized Medicine: Barriers to Biomarker Validation and Implementation**
    JAMA 2011;306(17):1914-1915. LD Fiore, LW D'Avolio

- **Next-Gen Phenotyping**
  - **Automated Concept-Level Information Extraction to Reduce the Need for Custom Software and Rules Development.**
    J Am Med Inform Assoc 2011 18(5): 607-13. LW D'Avolio L, T Nguyen, S Goryachev, LD Fiore
  - **Automated classification of psychotherapy note text: implications for quality assessment in PTSD care.**
    J Eval Clin Pract. 2012 Jun;18(3):698-701. B Shiner, LW D'Avolio, TM Nguyen, MH Zayed, BV Watts, LW Fiore

Contact

- Email:
  - ldavolio@gmail.com or leonard.davolio@va.gov
- Twitter:
  - @ldavolio
- Blog:
  - www.4thparadigm.org

- Online Registry of Biomedical Informatics Tools (ORBIT)
  - www.orbitproject.org or http://orbit.nlm.nih.gov/