

October 15, 2008

### I2b2 Academic Users Group UMassMed i2b2 Implementation Update Ralph J. Zottola, PhD

Fishers Lane Conference Center Rockville, MD

# UMassMed CTSA Planning

- A cross-functional team was assembled for the CTSA Biomedical Informatics planning group.
- Outcomes:
  - Need for both research faculty and core support for bio- and medical informatics
  - Need to invite the community to participate in clinical research
  - Critical need for access to clinical data
    - Clinical investigators complained that even with funding to pay for an analyst, it tool 6-9 months to get data.



I2b2 AUG: UMassMed Implementation Update

### **Outcomes: Need for Informatics Faculty**

- Created Program in Bioinformatics and Integrative Biology
  - Five new Faculty
  - Zhiping Weng, PhD recruited as Director
- Created Department of Quantitative Health Sciences
  - Academic home for biostatistics, epidemiology and medical informatics
  - Fifteen new Faculty



### **Outcome: Invite Patients to Clinical Research**

Establish the Conquering Diseases Biorepository (Dr. Craig Lilly, PI) which seeks to

- To provide a pathway that lets our patients support translational research in their diseases
- Welcome all of our patients to participate in clinical trials at UMass sites

We will do this by

- Creating a separate database of individuals who desire contact about participation in clinical research into their disease
- Creating a data repository of de-identified clinical information and linked patient derived samples that is available to qualified UMass researchers



### **Outcome: Need Access to Clinical Data**

- Adopted i2b2 framework to build out clinical research data repository (Dr. Ralph Zottola, PI)
- Emphasis on governance, legal, privacy and security issues
  - Need to address and not dismiss any concerns
- Leadership support critical
- IT led project
- Contracted with Recombinant Data Corporation as a planning and implementation partner.
- UMMHC licensed RDC's Quality Reporting System (developed at Partners)



## **I2b2 Implementation Goals**

- Community Participation
  - Consented bio-samples
  - Recruitment database
- Complete, current, accurate clinical data
- Link to clinical treatment (win/win/win)
- A extensible model for regional collaboration
  - Leverage NIH software from i2b2 and caBIG data integration tools
  - TIDE Trusted environment for data integration
  - Standard Operating Procedures



# **TIDE Architecture**

- The Trusted Independent Data Environment is the repository for all identified data
- Medical School to function as an "Honest Broker"
- Highly secure
  - Dedicated firewalls, IDS, two factor authentication
  - Limited number of users (currently 9 total)
  - No "internet access" all transfers via VPN secure FTP
  - Privacy training (CITI) and background checks for all people that have access
  - Regular audits of traffic and system usage
- SOPs for data management



## **TIDE Architecture**



I2b2 AUG: UMassMed Implementation Update



8

# **Creating a Clinical Data Repository**

- Data includes problems, medications, and labs
- Simple enterprise master index mechanism for data integration
- Extraction Transformation and Loading Step 1
  - Flat file extracts
  - Claims, billing, and scheduling systems
  - Data duplicates and badly formed data fixed
  - Data integrity among data sources is established using source dictionaries (e.g. ICD9)
- Result: a complete, quality tested clinical data repository



## **Clinical Data Integration**





### **Transforming Data for Specific Needs**

- ETL Step 2a: Research Repository
  - Map all data to i2b2 ontologies
  - Normalize all data types into facts (medications, problems, labs, etc.)
  - Anonymize by removing identifiers and generating nontraceable sequence numbers
- ETL Step 2b: Quality Data Mart
  - Create disease specific aggregations (i.e. diabetes, CAD)
  - Additional patient validation (panels)
  - Reporting brings data quality issues to the surface



## **Delivering Research Data**

- TIDE governance is integrated with IRB
- SOPs for data delivery and for re-integration of new research data into the repository (requires the adoption of proper protocols by researchers)
- Data can delivered to a private SQL server database that is accessed through a remote desktop
- Other data delivery processes can be created to accommodate approved protocols (e.g., breast cancer screening study)



I2b2 AUG: UMassMed Implementation Update

October 15, 2008

Ralph J. Zottola, PhD

### **Quality Reporting**



#### PHYSICIAN LAST, PHYSICIAN FIRST

#### DIABETES INTERVENTION REPORT

#### Generated on: 9/10/08

#### Report Key

| Latest date for each of the major data source |           | Test Result In-Range  |                      |  |  |  |
|---|-----------|---|----------------------|--|--|--|
| MEDITECH Labs                                 | 8/31/2008 | Test Result Value Out-of-Range                                  | X                    |  |  |  |
| Health Plan Pharmacy Claims                   | 8/16/2008 | Test Result Unknown   | 2                    |  |  |  |
| Health Plan Medical Claims                    | 8/30/2008 | Overdue Tests Shown with a Box around Most Recent Result        | Ė.                   |  |  |  |
| IDX Billing Data                              | 8/31/2008 | Coming DueTests Shown with a Triangle around Most Recent Result | $\overline{\Lambda}$ |  |  |  |
| IDX Scheduling Data                           | 5/6/2009  |   | _                    |  |  |  |

#### **Report Description**

The Diabetes Intervention Report contains at-a-glance information about your patients and their long-term diabetes care.

Past Three HbA1Cs holds the dates and results of the three most recent hemoglobin A1C tests. If HbA1c > 7.0 → Test Result Value Out of Range If last HbA1c > 6 months from run date → Test Overdue

#### Past Three LDLs holds the dates and results of the three most recent LDL tests.

If LDL > 100  $\rightarrow$  Test Result Value Out of Range If last LDL > 12 months from run date  $\rightarrow$  Test Overdue

Nephropathy Monitoring holds the dates and results of HEDIS Nephropathy Monitoring numerator satisfying events, which include evidence of known nephropathy, medication therapy or microalbuminuria testing in the past 12 months from run date.

KWN indicates evidence of known nephropathy

MED indicates the presence of numerator satisfying medication therapy (ACE/ARB) LAB indicates a microalbuminuria test

Next Scheduled Visit holds IDX Scheduling data, including visit date and provider. Status is the IDX Visit Type.



### I2b2 AUG: UMassMed Implementation Update

### **Quality Reporting**

### DIABETES INTERVENTION REPORT

| Provider:      | PHYSICIAN LAST, PHYSICIAN FIRST |
|----------------|---------------------------------|
| Provider Code: | 1083697841                      |
| Practice Code: | 15221 (Name of Clinic)          |

Patients in this report: 18 Page 2 of 7

| Patient Information   |                            | Past Three HbA1cs |                  |                  | Past Three LDLs  |                 |                  | Nephropathy<br>Monitoring | Next<br>Scheduled<br>Visit |
|---|----------------------------|-------------------|------------------|------------------|------------------|-----------------|------------------|---------------------------|----------------------------|
| LAST, FIRST<br>DOB: XX/XX/XXXX<br>SEX: M  | Status                     |                   | ~                | ~                | $\bigtriangleup$ | x               | x                | ?                         | RPE                        |
| MRN: 123123<br>Phone: 508-XXX-XXXX<br>Address: 123 Main St., Worcester,<br>MA 01605 | Result<br>Date<br>Provider | 6.6<br>9/24/2007  | 6.9<br>5/2/2007  | 6.8<br>1/25/2007 | 76<br>9/24/2007  | 105<br>5/2/2007 | 155<br>1/25/2007 |                           | 9/23/2008<br>ADLER         |
| LAST2, FIRST2<br>DOB: XX/XX/XXXX<br>SEX: M  | Status                     | x                 | x                | x                | ?                | ?               | ?                | *                         | EPB                        |
| MRN: 345345<br>Phone: 508-XXX-XXXX<br>Address: 345 Main St., Worcester,<br>MA 01605 | Result<br>Date<br>Provider | 8.9<br>4/8/2008   | 7.1<br>5/16/2006 | 7.3<br>9/16/2006 |                  |                 | -                | MED<br>7/7/2008           | 10/1/2008<br>ADLER         |



### **The Numbers**

- In production today, 18M facts on 175,000 patients
- Processing an additional 118 million lab facts and 25 million diagnoses
- Meditech DR project scheduled for completion by end of month
  - Clinical data repository will grow to 1.4M patients



I2b2 AUG: UMassMed Implementation Update

### **BioRepository Uses Workbench as Catalogue**



University of Massachusetts UMASS.Medical School

16