Data Centric Tools to Enable Clinical Research

Jomol Mathew, Ph.D.

Nov 16, 2016

Data Sciences & Technology
Data Sciences & Technology

Functions:
• Provide infrastructure, data and analytics support for basic, clinical and translational research
  • Systems & Tools for collection, integration and visualization of data
  • High performance Computing
  • Experienced staff consultation on data and tools
Tools for Clinical and Translational Research

Vision: Integrative Clinical & Biospecimen Data Ecosystem

Requires
• Tools
• Policies and procedures, especially for clinical data
• Collaboration with clinical system
New Tools

Patients

Patient self Reported/Device & Clinical Data

Research Data

Biospecimen Banking

Lab Archives

EPIC

Open Specimen

Data Lake

Clinical Decision Support

Informed Clinical Trials

Clinical trial management system

Volunteer Data Repository

Synergist TrinetX

Visualization & Analytics
The Data Powerhouse: Data Lake for Research (DLR)
Feasibility Assessment for Clinical Trial

Explore the patient population using de-identified data

e.g. Women With Endometriosis

• Do these patients exist at UMMS?
• What is the age and gender distribution?
• What other conditions exist in these patients?
• What medications do these patients take?
How do I use existing data to design the trial?

e.g. Botulinum Toxin for Pelvic Pain in Women With Endometriosis (NCT01553201)

• How does protocol design (exclusion/inclusion) impact recruitment?
• How can other sites selected for a multisite study?

INCLUSION CRITERIA:
• Female gender
• Age between 18 and 50
• History of endometriosis
• Persistent pelvic pain for at least 3 months
• Pelvic floor spasm
• Negative pregnancy test
• Willing to use reliable method of contraception for the month after botulinum toxin injection
• Willing and able to give informed consent
• Willing and able to comply with study requirements

EXCLUSION CRITERIA:
• Women with other causes of chronic pelvic pain including infectious, gastrointestinal, psychological disorders, fibromyalgia and chronic fatigue syndrome based on review of medical history within 1 year of first study visit*.
• Untreated severe cervical dysplasia or other gynecologic condition within the past year based on medical record review*.
• Significant abnormalities in the physical or laboratory examination including renal and liver function more than twice the normal range
• Hysterectomy and bilateral salpingo-oophorectomy
• Pregnancy
• Lactation
• Allergy to albumen or botulinum toxin
• Presence of antibodies to botulinum toxin or loss of response to previous injections for any indication
• A known neuromuscular junction disorder such as myasthenia gravis or Eaton-Lambert syndrome
• History of urinary or fecal incontinence
• Known pelvic prolapse
Informatics Tools to Assist:

- i2B2 & TrinetX
  - Both use i2b2 platform and stores de-identified patient data
  - Data elements include demographics, CPT, ICD, labs and medications

Sign up at: www.umassmed.edu/IT/CDP
TriNetX: Using existing data to help design Inclusion/Exclusion criteria

**INCLUSION CRITERIA:**
- Female gender
- Age between 18 and 50
- History of endometriosis

**EXCLUSION CRITERIA:**
- Fibromyalgia
- Cervical dysplasia
- Renal and liver function
### TriNetX: Comorbidities

#### Diagnoses of Patients in Cohort

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Patient Count</th>
<th>% of Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>N80-N98 Noninflammatory disorders of female genital tract</td>
<td>931</td>
<td>20%</td>
</tr>
<tr>
<td>200-213 Persons encountering health services for examinations</td>
<td>843</td>
<td>19%</td>
</tr>
<tr>
<td>R10-R19 Symptoms and signs involving the digestive system and abdomen</td>
<td>723</td>
<td>16%</td>
</tr>
<tr>
<td>Z77-Z99 Persons with potential health hazards related to family and personal history and certain conditions influencing health status</td>
<td>608</td>
<td>13%</td>
</tr>
<tr>
<td>R50-R69 General symptoms and signs</td>
<td>578</td>
<td>13%</td>
</tr>
<tr>
<td>Z30-Z39 Persons encountering health services in circumstances related to reproduction</td>
<td>498</td>
<td>11%</td>
</tr>
<tr>
<td>R00-R09 Symptoms and signs involving the circulatory and respiratory systems</td>
<td>480</td>
<td>11%</td>
</tr>
<tr>
<td>N70-N77 Inflammatory diseases of female pelvic organs</td>
<td>455</td>
<td>10%</td>
</tr>
<tr>
<td>D10-D36 Benign neoplasms, except benign neuroendocrine tumors</td>
<td>440</td>
<td>10%</td>
</tr>
<tr>
<td>Z20-Z28 Persons with potential health hazards related to communicable diseases</td>
<td>411</td>
<td>9%</td>
</tr>
<tr>
<td>G40-G47 Episodic and paroxysmal disorders</td>
<td>410</td>
<td>9%</td>
</tr>
<tr>
<td>J00-J06 Acute upper respiratory infections</td>
<td>404</td>
<td>9%</td>
</tr>
<tr>
<td>M50-M54 Other dorsopathies</td>
<td>390</td>
<td>9%</td>
</tr>
<tr>
<td>M20-M25 Other joint disorders</td>
<td>377</td>
<td>9%</td>
</tr>
<tr>
<td>M70-M79 Other soft tissue disorders</td>
<td>375</td>
<td>9%</td>
</tr>
<tr>
<td>N30-N39 Other diseases of the urinary system</td>
<td>362</td>
<td>8%</td>
</tr>
<tr>
<td>F40-F48 Anxiety, dissociative, stress-related, somatoform and other nonpsychotic mental disorders</td>
<td>359</td>
<td>8%</td>
</tr>
<tr>
<td>G69-G99 Other disorders of the nervous system</td>
<td>342</td>
<td>8%</td>
</tr>
<tr>
<td>F30-F39 Mood (affective) disorders</td>
<td>317</td>
<td>7%</td>
</tr>
<tr>
<td>K55-K64 Other diseases of intestines</td>
<td>297</td>
<td>7%</td>
</tr>
</tbody>
</table>
TriNetX: Concomitant Medications

547 unique medications

Medications of Patients in Cohort

- Central nervous system medications
- Respiratory tract medications
- Musculoskeletal medications
- Dermatological agents
- Antimicrobials
- Ophthalmic agents
- Gastrointestinal medications
- Hormones/synhetics/modifiers
- Nasal and throat agents, topical
- Genitourinary medications
- Rectal/Local
- Therapeutic nutrients/minerals/electrolytes
- Vitamins
- Cardiovascular medications
- Antihistamines
- Autonomic medications
- Ctic agents
- Antineoplastics
- Irrigation/Saline solutions
- Dental and oral agents, topical

Patient Count, % of Cohort
- 495 Central nervous system medications: 53%
- 458 Respiratory tract medications: 49%
- 445 Musculoskeletal medications: 47%
- 420 Dermatological agents: 42%
- 418 Antimicrobials: 44%
- 395 Ophthalmic agents: 42%
- 394 Gastrointestinal medications: 42%
- 381 Hormones/synhetics/modifiers: 41%
- 328 Nasal and throat agents, topical: 35%
- 275 Genitourinary medications: 29%
- 219 Rectal/Local: 23%
- 194 Therapeutic nutrients/minerals/electrolytes: 21%
- 190 Vitamins: 20%
- 189 Cardiovascular medications: 20%
- 182 Antihistamines: 19%
- 179 Autonomic medications: 19%
- 177 Ctic agents: 19%
- 109 Antineoplastics: 12%
- 96 Irrigation/Saline solutions: 10%
- 66 Dental and oral agents, topical: 7%
**TriNetX: Labs**

### Metabolic panel

<table>
<thead>
<tr>
<th>Test Code</th>
<th>Description</th>
<th>Mean ± SD</th>
<th>Min</th>
<th>Max</th>
<th>Patient Counts, % of Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>9029</td>
<td>Sodium (Moles/volume) in Serum, Plasma or Blood</td>
<td>137.83 ± 2.24</td>
<td>130</td>
<td>147</td>
<td>557 (59%)</td>
</tr>
<tr>
<td>9028</td>
<td>Potassium (Moles/volume) in Serum, Plasma or Blood</td>
<td>4.11 ± 0.39</td>
<td>2.7</td>
<td>6.2</td>
<td>557 (59%)</td>
</tr>
<tr>
<td>9023</td>
<td>Chloride (Moles/volume) in Serum, Plasma or Blood</td>
<td>104.71 ± 2.85</td>
<td>93</td>
<td>118</td>
<td>556 (59%)</td>
</tr>
<tr>
<td>9021</td>
<td>Bicarbonate (Moles/volume) in Serum, Plasma or Blood</td>
<td>26.16 ± 2.58</td>
<td>14</td>
<td>33</td>
<td>553 (59%)</td>
</tr>
<tr>
<td>9030</td>
<td>Urea nitrogen (Moles/volume) in Serum, Plasma or Blood</td>
<td>1.21 ± 7.03</td>
<td>0.28</td>
<td>121</td>
<td>582 (62%)</td>
</tr>
<tr>
<td>9024</td>
<td>Creatinine [Mass/volume] in Serum, Plasma or Blood</td>
<td>99.41 ± 32.15</td>
<td>46</td>
<td>424</td>
<td>593 (63%)</td>
</tr>
<tr>
<td>9025</td>
<td>Glucose [Mass/volume] in Serum, Plasma or Blood</td>
<td>9.29 ± 0.55</td>
<td>7.5</td>
<td>14</td>
<td>533 (57%)</td>
</tr>
<tr>
<td>9022</td>
<td>Calcium [Mass/volume] in Serum, Plasma or Blood</td>
<td>2.18 ± 0.97</td>
<td>1.4</td>
<td>9.8</td>
<td>131 (14%)</td>
</tr>
<tr>
<td>9026</td>
<td>Magnesium [Mass/volume] in Serum, Plasma or Blood</td>
<td>3.56 ± 0.82</td>
<td>1.3</td>
<td>8</td>
<td>109 (12%)</td>
</tr>
</tbody>
</table>

### Complete blood count

<table>
<thead>
<tr>
<th>Test Code</th>
<th>Description</th>
<th>Mean ± SD</th>
<th>Min</th>
<th>Max</th>
<th>Patient Counts, % of Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>9012</td>
<td>Erythrocytes [#/volume] in Blood</td>
<td>8.02 ± 2.85</td>
<td>3</td>
<td>22.9</td>
<td>660 (70%)</td>
</tr>
<tr>
<td>9015</td>
<td>Leukocytes [#/volume] in Blood</td>
<td>12.73 ± 1.57</td>
<td>5.8</td>
<td>17.1</td>
<td>666 (71%)</td>
</tr>
<tr>
<td>9014</td>
<td>Hemoglobin [Mass/volume] in Blood</td>
<td>37.17 ± 4.99</td>
<td>19.3</td>
<td>49.1</td>
<td>679 (72%)</td>
</tr>
<tr>
<td>9020</td>
<td>Platelets [#/volume] in Blood</td>
<td>254.73 ± 66.07</td>
<td>64</td>
<td>839</td>
<td>660 (70%)</td>
</tr>
<tr>
<td>9008</td>
<td>Erythrocyte distribution width (Ratio)</td>
<td>88.75 ± 6.11</td>
<td>58.1</td>
<td>109.8</td>
<td>668 (71%)</td>
</tr>
<tr>
<td>9009</td>
<td>Erythrocyte mean corpuscular volume [Entitic volume]</td>
<td>29.85 ± 2.54</td>
<td>16.5</td>
<td>36.5</td>
<td>668 (71%)</td>
</tr>
<tr>
<td>9011</td>
<td>Erythrocyte mean corpuscular hemoglobin [Entitic mass]</td>
<td>33.60 ± 0.97</td>
<td>28.3</td>
<td>35.9</td>
<td>670 (71%)</td>
</tr>
</tbody>
</table>

### Additional tests

- 9019: Platelet mean volume [Entitic volume] in Blood
- 9018: Neutrophils [#/volume] in Blood
- 9016: Lymphocytes/100 leukocytes in Blood
I would like to do a retrospective study using all data (including PHI). Can I get data on these patients?

* e.g. Women With Endometriosis
  * Detailed Assessment of comorbidities
  * Follow the cohort – long term QOL, development of other diseases
  * Need to link to specimens
Note from the IRB

If you are requesting identifiable data (PHI/PII) from the Clinical Data Portal, you must do so under:

– A HIPAA Waiver and/or HIPAA Authorization
– Other appropriate documentation
Detailed Data access

How:
• www.umassmed.edu/IT/CDP

Who:
• Faculty: Instructor or above
• Any member of a research team
• Administrators and staff at UMMHC or UMMS
Note from the IRB

IRB Documents

Required to Obtain PHI/PII Data for Research Purposes

Must be Consistent!
Streamlined Data Access Policies @ UMMS

Key points:
- De-identified or Aggregate data: No IRB approval required
- Protected Health information (PHI): IRB approval required
- If in doubt, ask the IRB
Enabling Research Participant Recruitment

• Recruit volunteers for your studies via Volunteer Registry
• E-consent pilot in process for Volunteer Registry

Expanding via
• Social media
• Special population resource center
• Direct to patient tools
• Recruiting via EHR once EPIC in place
Analytical Capabilities: Geographic Pattern Finding: GIS
Basic & Translational Research Tools

• Biospecimen Banking: OpenSpecimen
• Electronic Lab Notebook: Lab Archives
• Search & Share Data: Synergist
• MGHPC & a myriad of genomics tools
Biospecimen Banking: OpenSpecimen

**Single Shop for Biospecimens**
- Consent, collect & barcode
- Create derivatives & keep lineage
- Search & find
- Scan & distribute
- Link to clinical data in Data Lake & facilitate query and request of biospecimens from central biobanks (blood, tumor, microbiome)
OpenSpecimen: Helps Keep Specimen Lineage

<table>
<thead>
<tr>
<th>Parent Specimen</th>
<th>Available Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>036V15000480.12</td>
<td>8.600000e+6 Cells</td>
</tr>
</tbody>
</table>

**Create Derived Specimen**

- **Specimen Type**: PBMC
- **Anatomic Site**: Blood
- **Pathology Status**: Not Specified
- **Created On**: 09-12-2016
- **Freeze/Thaw Cycles**: 1

Additional fields:
- **Type**, **Label**, **Quantity**, **Concentration**, **Location**, **Comments**, **Media Volume**, **Cell Viability (%)**

Options:
- **Do you want to close parent specimen?**
- **Create**
- **Discard**
OpenSpecimen: Status

• Piloted with Dr. Luzuriaga Lab – went live on 09/12/2016
• Next Wave: Central banks

• Guiding Principles:
  – Researchers **do not** need IRB approval to obtain de-identified samples
  – Identified specimens can be provided after IRB approval
Collect & Manage Research Data: Lab Archives

- Enables easy access to data between lab members and collaborators
- Supports secure data trail (necessary for commercialization)
Search & Share Research Data: Synergist

“Amazon” of Research data
- Catalog & Share Experimental Metadata
- Search and Discover Data & Collaborate
- Connect & Gain Insights
- Publish & Submit Data to External Data Banks
Note from the IRB

If you will utilize the research tools described, it very important to:

- Think about and plan *(ahead)* for data sharing.
- Consider the consent form: Does it allow me to share in all the ways that I want to, now and in the future?
- Consider Genomic Data: Will I submit data to dbGaP?
  - It is necessary to obtain an Institutional Certification to share data with federal databases (e.g. dbGaP).
    » Adherence to the NIH Genomic Data Sharing (GDS) Policy is required.
    » The [NIH Points to Consider for IRBs and Institutions PDF](mailto:umms-otm@umassmed.edu), outlines the genomic data sharing risks that are to be included in the consent form.
- Consider applicable agreements:
  - What are the limits in my Data Use Agreement(s), Confidentiality agreements?
  - Do I have all the necessary agreements in place to protect intellectual property?
    - Contact the Office of Technology Management for guidance ([UMMS-OTM@umassmed.edu](mailto:umms-otm@umassmed.edu))
Integrative Data Ecosystem

Provide Capabilities to Identify Patterns/Classify/Compare/Contrast Patients using Integrative Clinical and Molecular/Genomic Data

- Match patients to clinical trials
- Enable better and efficient design of clinical trials
- Identify meaningful points for intervention for ‘continuous monitoring’ scenarios
- Predict outcomes & Suggest treatment options

Translation Enablers
Connect Research & Clinical Data

User queries for clinical information

Clinical Data Repository

Identified Data

TriNetx

Aggregate/De-identified Data

De-identified Patient Set

Honest Broker System

Corresponding Patient Set generated

Data Catalog

Identified Patient Set

Identified Data
Oncore CTMS

- Provides a searchable view of accurate and up-to-date open study information
- Enhance safety through EMR integration
- Standardized collection and reporting of clinical trial information
- Consistent tracking of safety data and staff credentials
- Email alerts and reminders
- Improved communication across team and sites
- Automation support (subject calendars & re-consenting, sponsor invoicing)
- EMR integration for Patient Safety
- Grants reporting

Researchers and Research Staff

Department / Admin.

Regulatory

Public

-Visibility to clinical research portfolio at different levels within the enterprise
- Resource allocation
- Effectively manage staff workload
  - Streamline budget negotiations
  - Billing compliance

Provides a searchable view of accurate and up-to-date open study information
- Enhance safety through EMR integration
OnCore: Visibility for Study Team & Leadership

Chairs and Administrators: Departmental view of
• Planned trials
• Open trials
• Enrollment
• Accounts billable/receivable

Custom Reports can be created upon request
### Timeline for OnCore

**Dry Run Begins**: 08/08/16

**OnCore Goes Live**
- **10/21/16**

**Onboarding Begins**
- **10/30/16**

**Onboarding Ends**
- **12/31/16**

**Onboarding Begins**
- **01/23/17**

**Onboarding Ends**
- **04/24/17**

**Onboarding Begins**
- **04/10/17**

**Onboarding Ends**
- **07/01/17**

**Epic Goes Live**: 10/01/17

**Onboarding Ends**: 12/31/16

### Pilot Teams Series 1 Series 2 Series 3

<table>
<thead>
<tr>
<th>Pilot Teams</th>
<th>Series 1</th>
<th>Series 2</th>
<th>Series 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurology</td>
<td>Clinical Research Center (CRC)</td>
<td>Emergency Medicine</td>
<td>Orthopedic</td>
</tr>
<tr>
<td>GI</td>
<td>Anesthesiology</td>
<td>Psychiatry</td>
<td>QHS</td>
</tr>
<tr>
<td>Cardiovascular Medicine</td>
<td>Remaining divisions* within Medicine</td>
<td>OB/GYN</td>
<td>Radiology*</td>
</tr>
<tr>
<td>Cancer Research</td>
<td>Surgery*</td>
<td>Diabetes Ctr. of Excellence</td>
<td>Pathology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pediatric*</td>
<td></td>
</tr>
</tbody>
</table>

* Tentative
Draft Policy for CTMS Use

Scope

All clinical research (as defined by the NIH) that meets any of the following criteria will be entered into CTMS:

- Purchases or uses a service from UMass Memorial Health Care (UMMHC) or any of its affiliates, including UMass Memorial Medical Group;
- Uses the CCTS Clinical Research Center;
- Has milestone based billing or payments;
- Will flag patients in Epic or use Epic for recruitment

Beginning July 1, 2017, entry into and use of the CTMS will be a prerequisite for study initiation, account set-up and for ongoing financial management of the study.

Exceptions granted on a case by case basis
Thanks to....

- **Advisory Committee** (K. Luzuriaga, C. Kiefe, S. Corvera, M. Koziel, N. Hafer, G. Wolf, J. Mathew)
- **EPIC Research Integrated Work Group** (T. Houston & M. Koziel, J. Mathew, D. Amante, B. Barton, E. Boudreaux, S. Cutrona, P. Franklin, and a lot of people from Epic and UMMHC)
Questions?

On the web:  
http://www.umassmed.edu/it/cdp/

By email:  
Jomol.Mathew@umassmed.edu  
Margaret.Koziel@umassmed.edu  
ClinicalDataPortal@umassmed.edu