



Data Centric Tools to Enable Clinical Research

Jomol Mathew, Ph.D.

Nov 16, 2016

Data Sciences & Technology

Data Sciences & Technology

Functions:

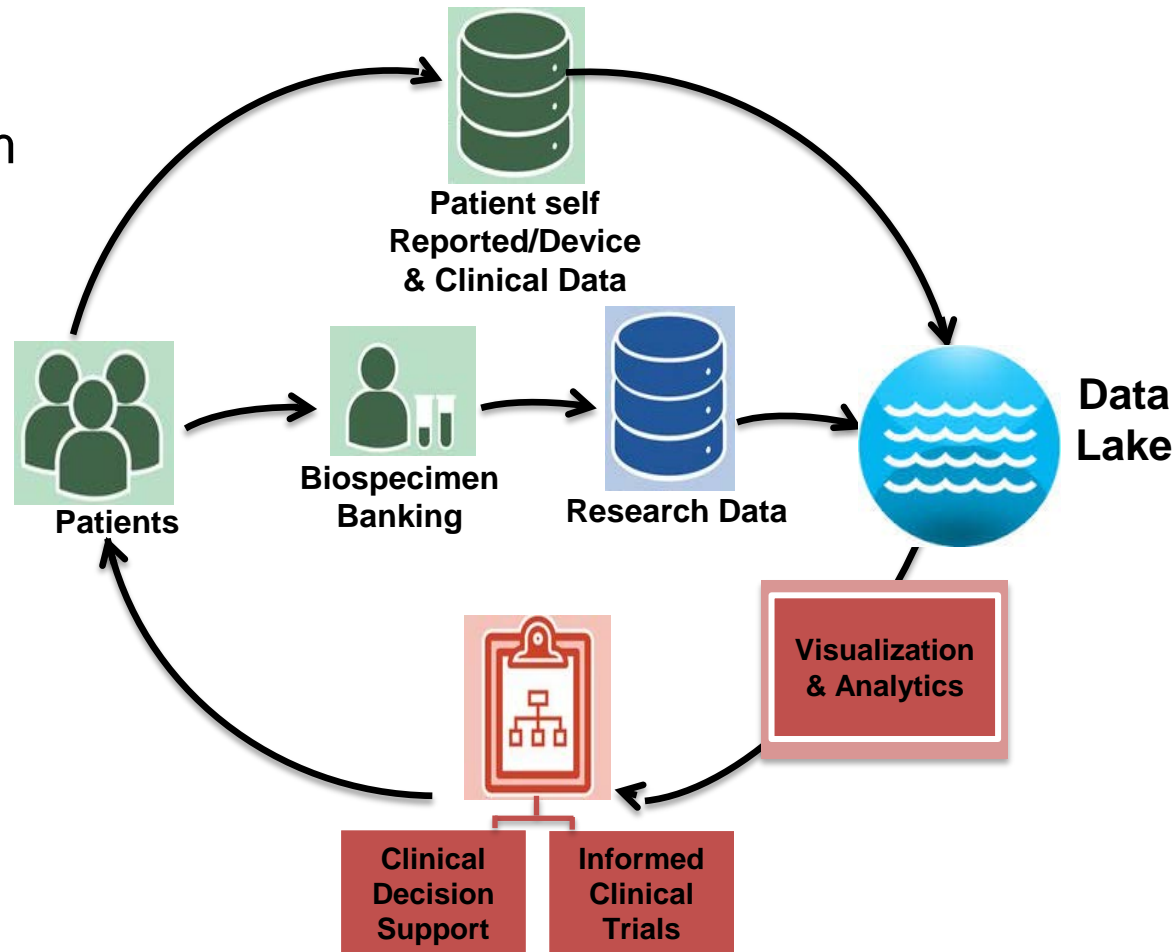
- Provide infrastructure, data and analytics support for basic, clinical and translational research
 - Systems & Tools for collection, integration and visualization of data
 - High performance Computing
 - Experienced staff consultation on data and tools

Tools for Clinical and Translational Research

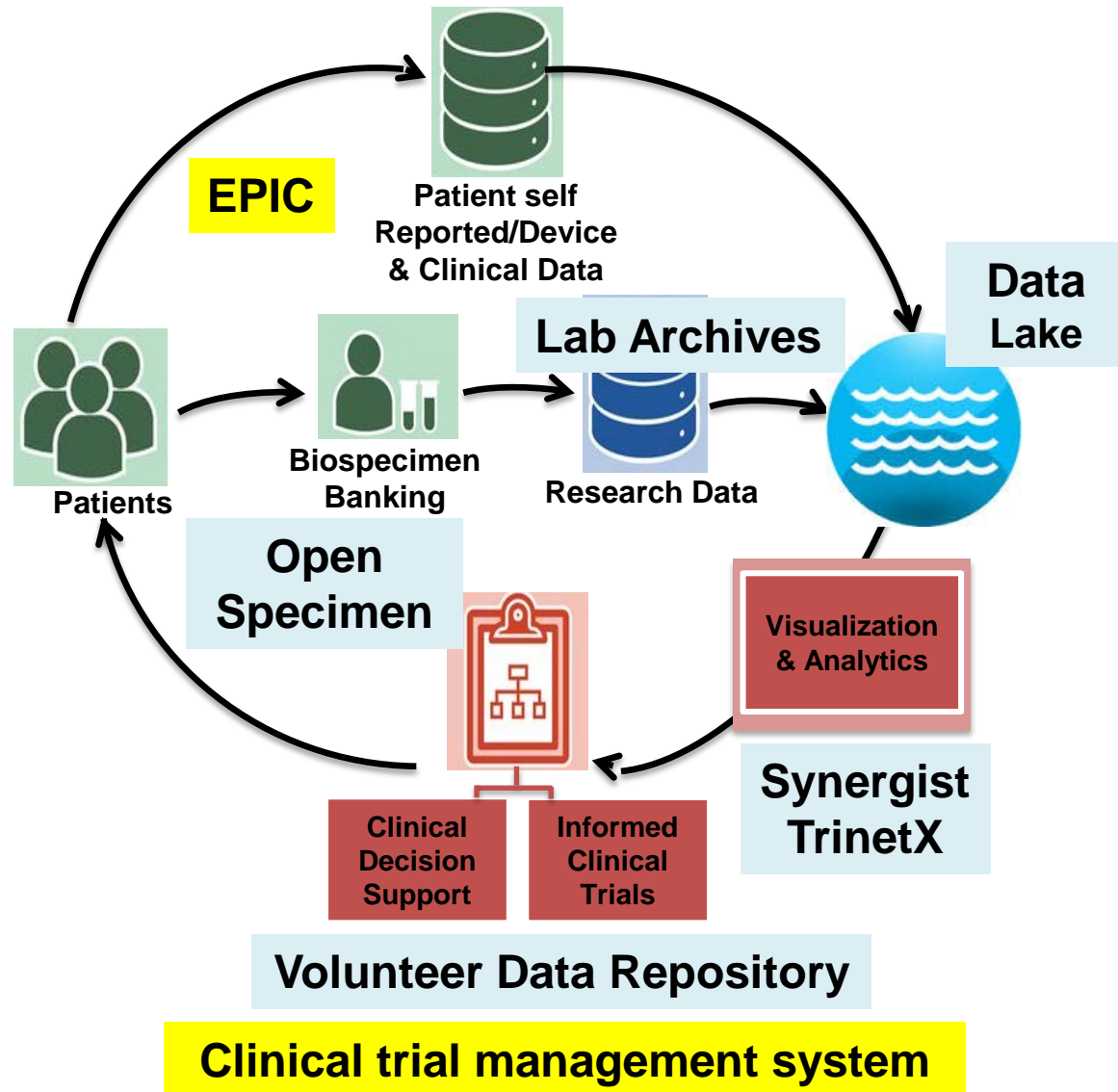
Vision: Integrative Clinical & Biospecimen Data Ecosystem

Requires

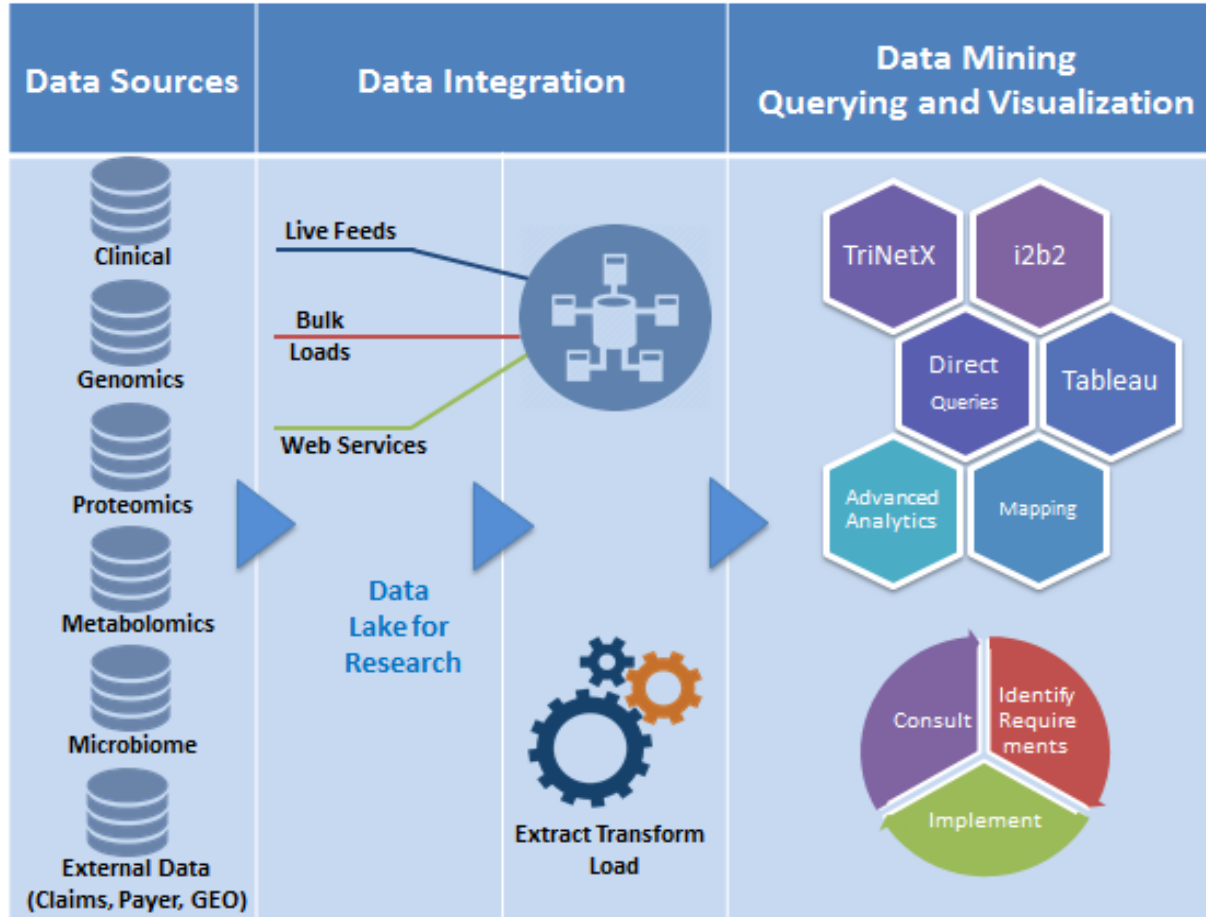
- Tools
- Policies and procedures, especially for clinical data
- Collaboration with clinical system



New Tools



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

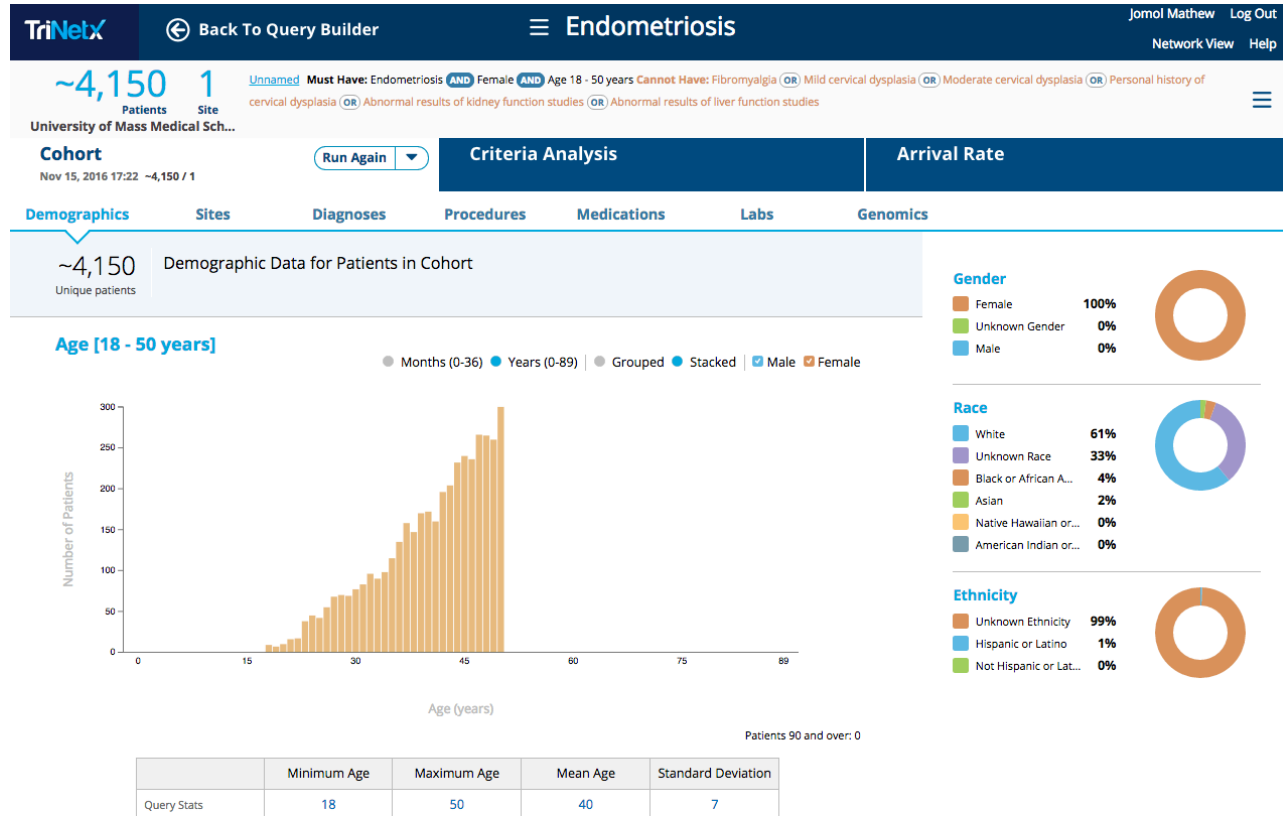
TriNetX: Using existing data to help design Inclusion/Exclusion criteria

INCLUSION CRITERIA:

- Female gender
- Age between 18 and 50
- History of endometriosis

EXCLUSION CRITERIA:

- Fibromyalgia
- Cervical dysplasia
- Renal and liver function



TriNetX: Comorbidities

TriNetX
Back To Query Builder
Endometriosis
Jomol Mathew Log Out

~4,150 Patients
1 Site
Unnamed Must Have: Endometriosis AND Female AND Age 18 - 50 years Cannot Have: Fibromyalgia OR Mild cervical dysplasia OR Moderate cervical dysplasia OR Personal history of cervical dysplasia OR Abnormal results of kidney function studies OR Abnormal results of liver function studies

University of Mass Medical Sch...
Run Again
Criteria Analysis
Arrival Rate

Nov 15, 2016 17:22 ~4,150 / 1

Demographics
Sites
Diagnoses
Procedures
Medications
Labs
Genomics

5,621
Unique diagnoses

Diagnoses of Patients in Cohort

Click on a term to add or remove it from the query

These results are for ~940 out of ~4,150 patients

	Patient Count, % of Cohort
▶ N80-N98 Noninflammatory disorders of female genital tract	931 ██████████ 99%
▶ Z00-Z13 Persons encountering health services for examinations	843 ██████████ 90%
▶ R10-R19 Symptoms and signs involving the digestive system and abdomen	723 ██████████ 77%
▶ Z77-Z99 Persons with potential health hazards related to family and personal history and certain conditions influencing health status	608 ██████████ 65%
▶ R50-R69 General symptoms and signs	578 ██████████ 61%
▶ Z30-Z39 Persons encountering health services in circumstances related to reproduction	498 ██████████ 53%
▶ R00-R09 Symptoms and signs involving the circulatory and respiratory systems	480 ██████████ 51%
▶ N70-N77 Inflammatory diseases of female pelvic organs	455 ██████████ 48%
▶ D10-D36 Benign neoplasms, except benign neuroendocrine tumors	440 ██████████ 47%
▶ Z20-Z28 Persons with potential health hazards related to communicable diseases	411 ██████████ 44%
▶ G40-G47 Episodic and paroxysmal disorders	410 ██████████ 44%
▶ J00-J06 Acute upper respiratory infections	404 ██████████ 43%
▶ M50-M54 Other dorsopathies	390 ██████████ 41%
▶ M20-M25 Other joint disorders	377 ██████████ 40%
▶ M70-M79 Other soft tissue disorders	375 ██████████ 40%
▶ N30-N39 Other diseases of the urinary system	362 ██████████ 39%
▶ F40-F48 Anxiety, dissociative, stress-related, somatoform and other nonpsychotic mental disorders	359 ██████████ 38%
▶ G89-G99 Other disorders of the nervous system	342 ██████████ 36%
▶ F30-F39 Mood [affective] disorders	317 ██████████ 34%
▶ K55-K64 Other diseases of intestines	297 ██████████ 32%

0/0

Show

- Last 3 Months
- Last 6 Months
- Last 12 Months
- Last 24 Months
- Any Time

- All
- Acute **44%**
- Chronic **25%**

TriNetX: Concomitant Medications

TriNetX
Jomol Mathew [Log Out](#)

[Back To Query Builder](#)
[Endometriosis](#)

[Network View](#) [Help](#)

~4,150 Patients **1** Site

University of Mass Medical Sch...

Cohort
Nov 15, 2016 17:22 ~4,150 / 1

[Run Again](#)

Criteria Analysis | **Arrival Rate**

[Demographics](#) | [Sites](#) | [Diagnoses](#) | [Procedures](#) | [Medications](#) | [Labs](#) | [Genomics](#)

547 Unique medications | Medications of Patients in Cohort
Click on a term to add or remove it from the query

These results are for ~940 out of ~4,150 patients

	Patient Count, % of Cohort
▶ N0000029132 Central nervous system medications	495 ██████████ 53%
▶ N0000029237 Respiratory tract medications	458 ██████████ 49%
▶ N0000029206 Musculoskeletal medications	445 ██████████ 47%
▶ N0000029154 Dermatological agents	426 ██████████ 45%
▶ N0000029074 Antimicrobials	418 ██████████ 44%
▶ N0000029216 Ophthalmic agents	395 ██████████ 42%
▶ N0000029168 Gastrointestinal medications	394 ██████████ 42%
▶ N0000029177 Hormones/synthetics/modifiers	381 ██████████ 41%
▶ N0000029212 Nasal and throat agents,topical	328 ██████████ 35%
▶ N0000029260 Genitourinary medications	275 ██████████ 29%
▶ N0000029248 Rectal,local	219 ██████████ 23%
▶ N0000029253 Therapeutic nutrients/minerals/electrolytes	194 ██████████ 21%
▶ N0000029267 Vitamins	190 ██████████ 20%
▶ N0000029116 Cardiovascular medications	189 ██████████ 20%
▶ N0000029071 Antihistamines	182 ██████████ 19%
▶ N0000029103 Autonomic medications	179 ██████████ 19%
▶ N0000029231 Otic agents	177 ██████████ 19%
▶ N0000029091 Antineoplastics	109 ██████████ 12%
▶ N0000029202 Irrigation/dialysis solutions	96 ██████████ 10%
▶ N0000029225 Dental and oral agents,topical	66 ██████████ 7%

[Show All](#)

Search: [+](#) [-](#)

0/0

Show

- Last 3 Months
- Last 6 Months
- Last 12 Months
- Last 24 Months
- Any Time

TriNetX: Labs

TriNetX
Back To Query Builder
Endometriosis
Jomol Mathew Log Out
Network View Help

~4,150 Patients
1 Site
Unnamed Must Have: Endometriosis AND Female AND Age 18 - 50 years Cannot Have: Fibromyalgia OR Mild cervical dysplasia OR Moderate cervical dysplasia OR Personal history of cervical dysplasia OR Abnormal results of kidney function studies OR Abnormal results of liver function studies

University of Mass Medical Sch...
Cohort
Run Again
Criteria Analysis
Arrival Rate

Nov 15, 2016 17:22 ~4,150 / 1

Demographics Sites Diagnoses Procedures Medications Labs Genomics

166 Unique labs
Most Recent Lab Values for Patients in the Cohort
Click on a term to add or remove it from the query

These results are for ~940 out of ~4,150 patients

		Mean ± SD	Min	Max	Patient Counts, % of Cohort	
Metabolic panel						
9029	Sodium [Moles/volume] in Serum, Plasma or Blood	137.83 ± 2.24	130	147	557	59%
9028	Potassium [Moles/volume] in Serum, Plasma or Blood	4.11 ± 0.39	2.7	6.2	557	59%
9023	Chloride [Moles/volume] in Serum, Plasma or Blood	104.71 ± 2.85	93	118	556	59%
9021	Bicarbonate [Moles/volume] in Serum, Plasma or Blood	26.16 ± 2.58	14	33	553	59%
9030	Urea nitrogen [Mass/volume] in Serum, Plasma or Blood				0	0%
9024	Creatinine [Mass/volume] in Serum, Plasma or Blood	1.21 ± 7.03	0.28	121	582	62%
9025	Glucose [Mass/volume] in Serum, Plasma or Blood	99.41 ± 32.15	46	424	593	63%
9022	Calcium [Mass/volume] in Serum, Plasma or Blood	9.29 ± 0.55	7.5	14	533	57%
9026	Magnesium [Mass/volume] in Serum, Plasma or Blood	2.16 ± 0.97	1.4	9.8	131	14%
9027	Phosphate [Mass/volume] in Serum, Plasma or Blood	3.56 ± 0.82	1.3	8	109	12%
Complete blood count						
9012	Erythrocytes [# /volume] in Blood				0	0%
9015	Leukocytes [# /volume] in Blood	8.02 ± 2.85	3	22.9	660	70%
9014	Hemoglobin [Mass/volume] in Blood	12.73 ± 1.57	5.8	17.1	666	71%
9013	Hematocrit [Volume Fraction] of Blood	37.17 ± 4.99	19.3	49.1	679	72%
9020	Platelets [# /volume] in Blood	254.23 ± 66.07	64	839	660	70%
9008	Erythrocyte distribution width [Ratio]				0	0%
9011	Erythrocyte mean corpuscular volume [Entitic volume]	88.75 ± 6.11	58.1	109.8	668	71%
9009	Erythrocyte mean corpuscular hemoglobin [Entitic mass]	29.85 ± 2.54	16.5	36.5	668	71%
9010	Erythrocyte mean corpuscular hemoglobin concentration [Mass/volume]	33.60 ± 0.97	28.3	35.9	670	71%
9019	Platelet mean volume [Entitic volume] in Blood	8.71 ± 1.07	6	13.7	668	71%
9018	Neutrophils [# /volume] in Blood	5.54 ± 2.82	1.4	21.16	561	60%
9016	Lymphocytes/100 leukocytes in Blood				0	0%

find 0/0

Show

- Last 3 Months
 - Last 6 Months
 - Last 12 Months
 - Last 24 Months
 - Any Time
-
- All Labs
 - Common Labs

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

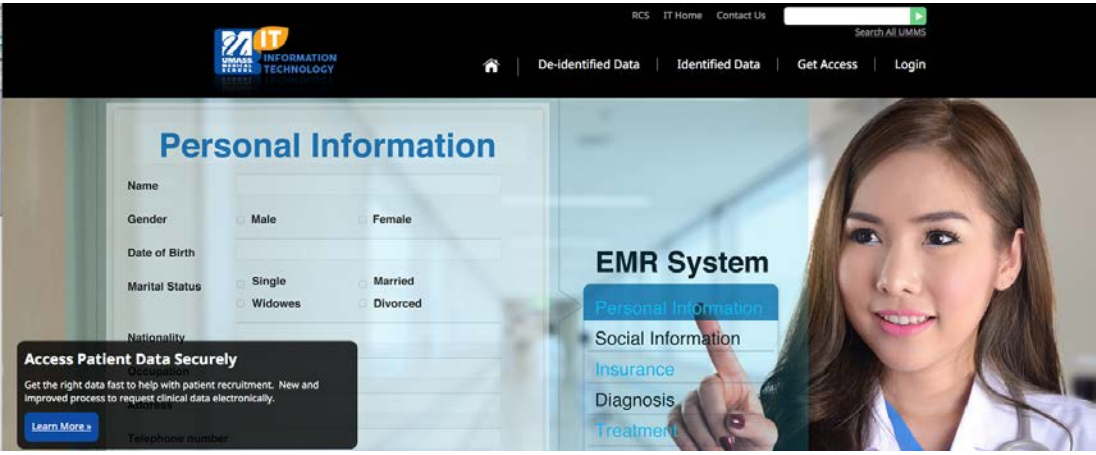


Note from the IRB

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

Study description including data criteria and special instructions
* must provide value

Request data from these clinical data resources:
* must provide value

- Soarian
- AllScripts
- IDX
- Payer Systems
- All of the above

Are you requesting any Protected Health Information (PHI) or Personally identifiable information (PII)?
* must provide value

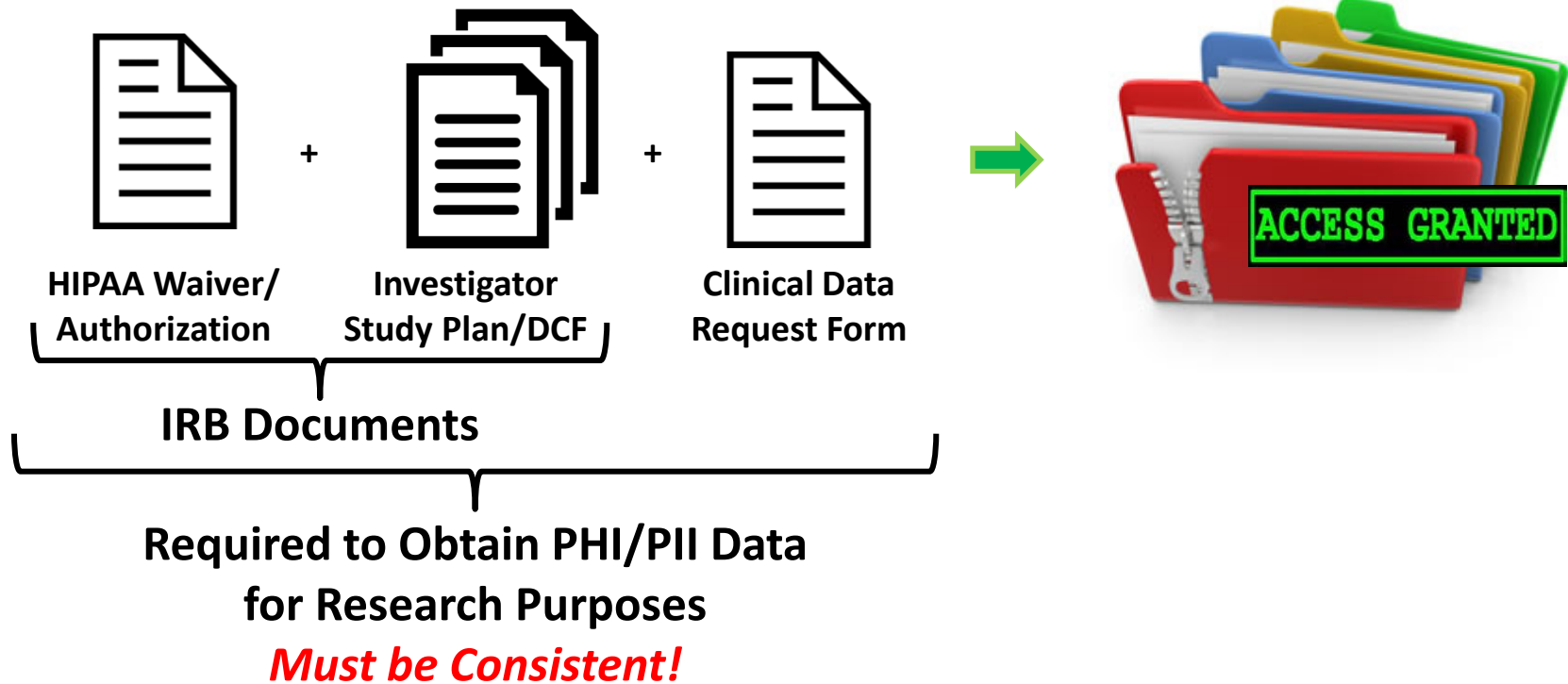
- Yes
- No

Request Date Range

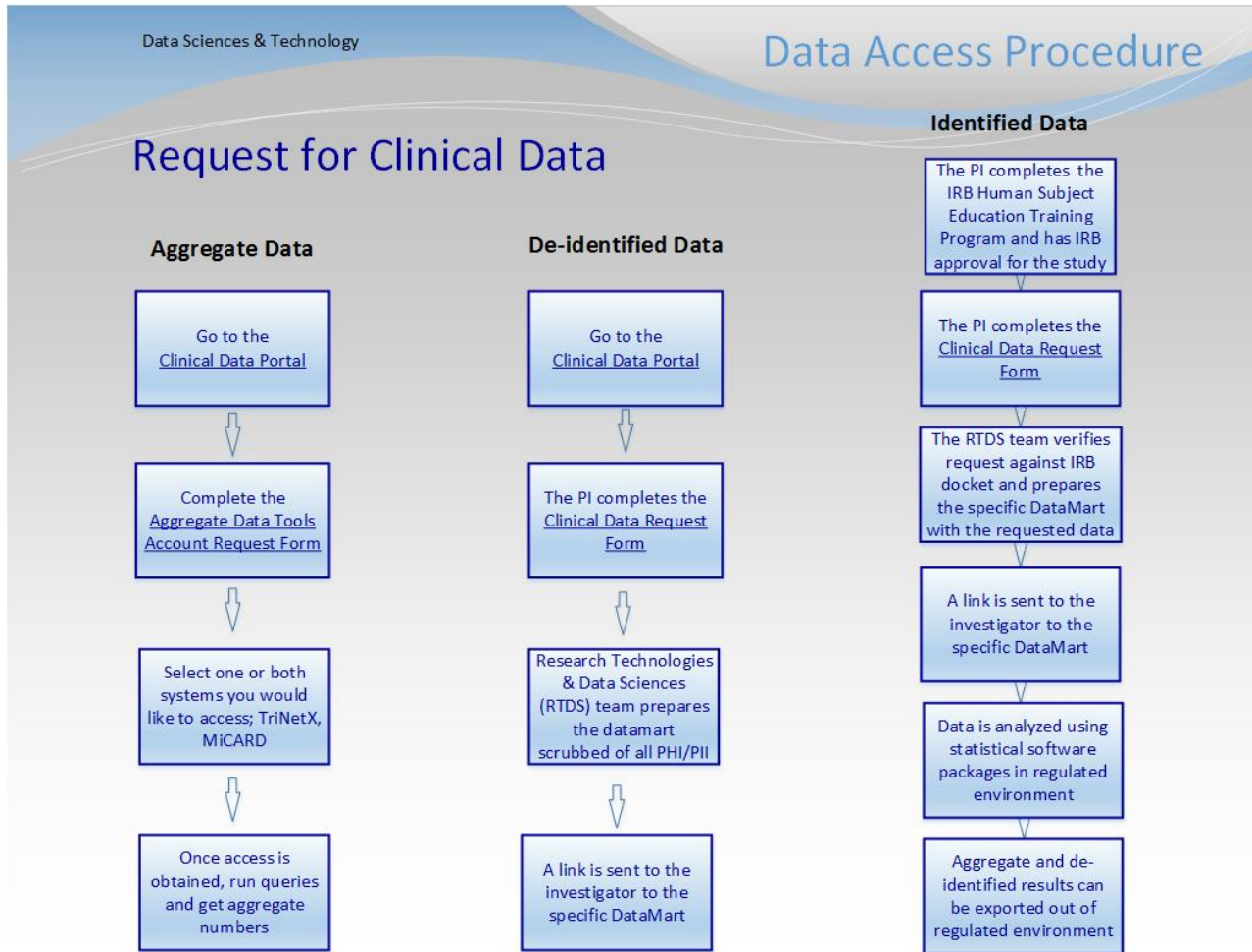
Begin Date 3.1 Today M-D-Y
(MM-DD-YYYY)



Note from the IRB



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

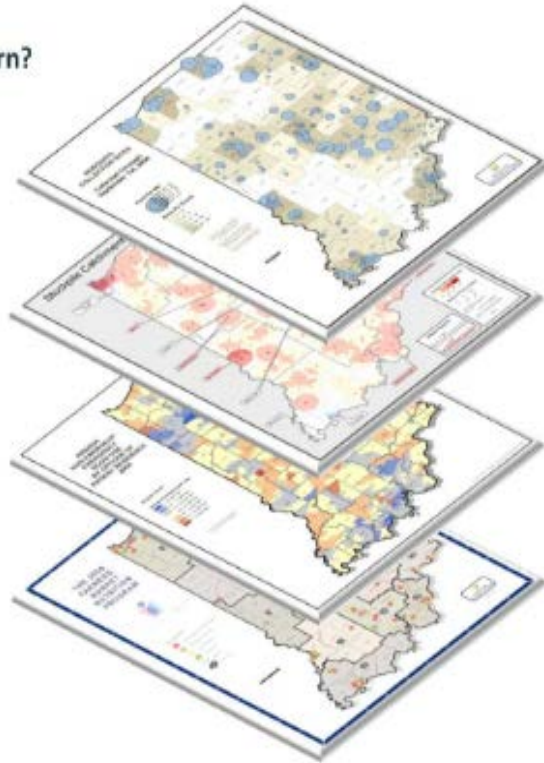
- Social media
- Special population resource center
- Direct to patient tools
- Recruiting via EHR once EPIC in place

Analytical Capabilities: Geographic Pattern Finding: GIS

How much?

Is there a pattern?

Where?



Who?

How far?

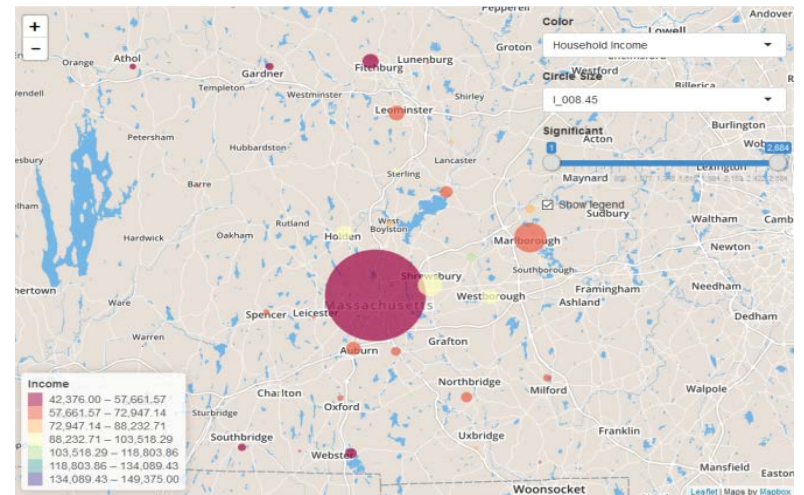
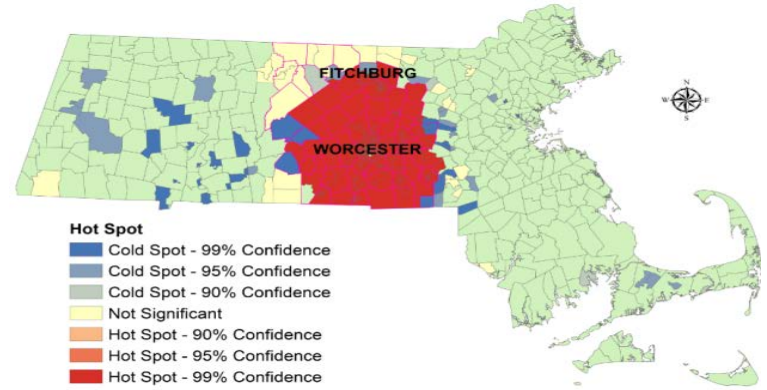
Why?

Likelihood?

How many?

Which one?

Is it a cluster?



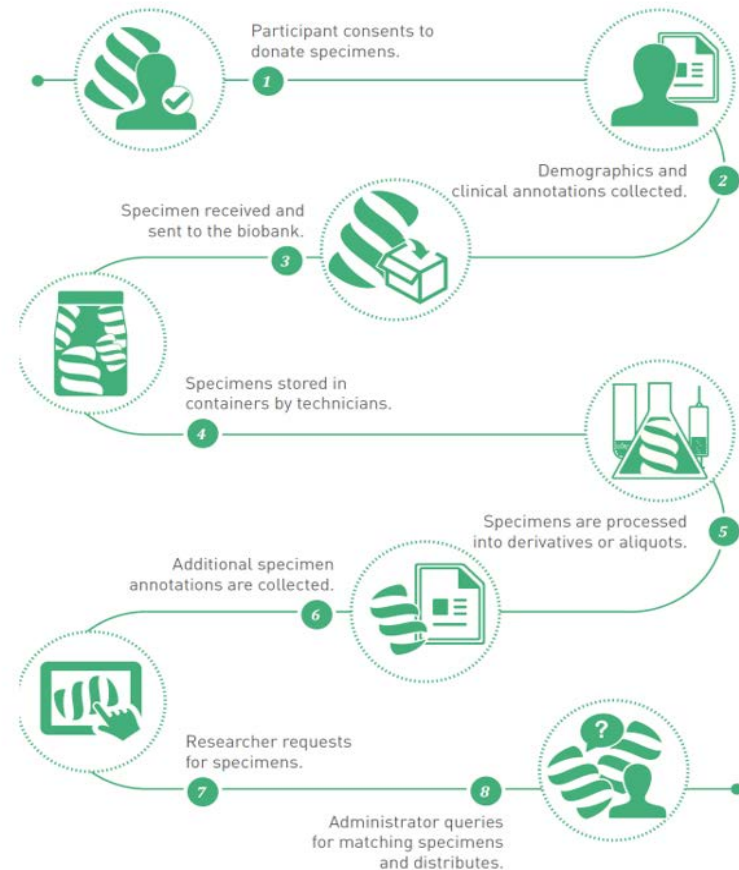
Basic & Translational Research Tools

- Biospecimen Banking: OpenSpecimen
- Electronic Lab Notebook: Lab Archives
- Search & Share Data: Synergist
- MGHPCC & 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



... / T14: Visit 3 - Week 2 / 036V15000480.12

Create Derived Specimen

Parent Specimen	036V15000480.12	Available Quantity	8.600000e+6 Cells
Specimen Type	PBMC	Anatomic Site	Blood
Type	<input type="text" value="Type"/>		
Label	<input type="text" value="Label"/>		
Quantity	<input type="text" value="Quantity"/>		
Concentration	<input type="text" value="Concentration"/>		
Location	<input type="text" value="Container"/>	<input type="text" value="Row"/>	<input type="text" value="Column"/>
Pathology Status	<input type="text" value="Not Specified"/>		
Created On	<input type="text" value="09-12-2016"/>	<input type="text" value="19"/>	<input type="text" value="45"/>
Freeze/Thaw Cycles	<input type="text" value="1"/>		
	<input checked="" type="checkbox"/> Increment parent specimen's Freeze/Thaw cycles?		
Comments	<input type="text" value="Comments"/>		
Media Volume	<input type="text"/>		
Cell Viability (%)	<input type="text"/>		

Do you want to close parent specimen?

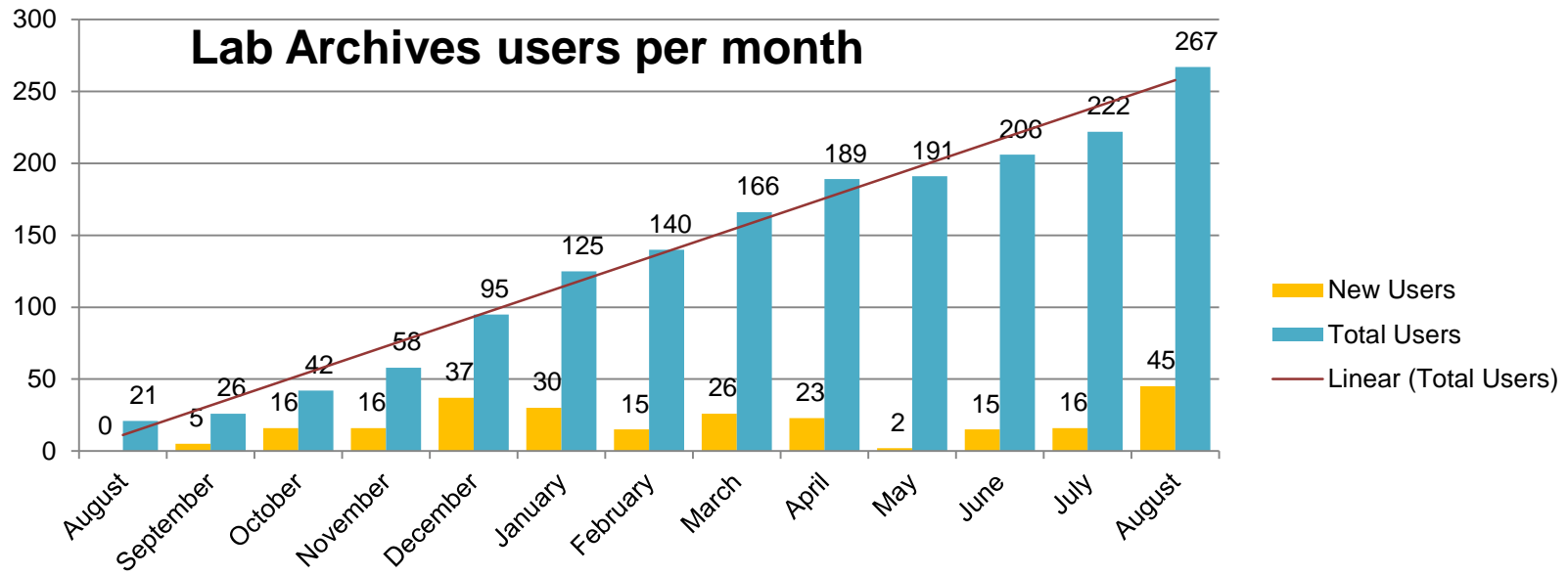
Discard

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 & Mangle Research Data: Lab Archives

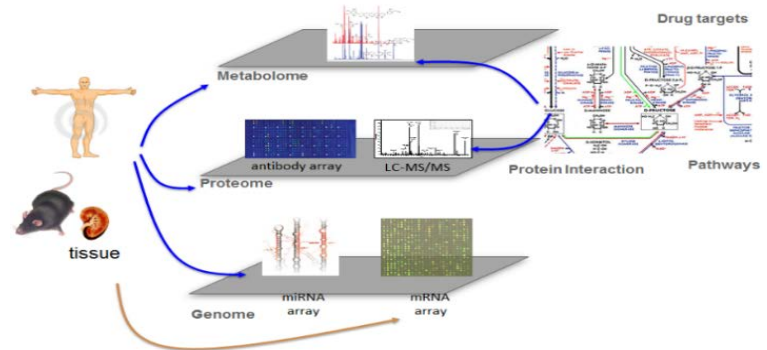
- Enables easy access to data between lab members and collaborators
- Supports secure data trail (necessary for commercialization)



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



The screenshot shows the Synergist web interface. The top navigation bar includes 'Home', 'Help', 'Contact', 'Login', 'Export', 'Dashboard', and a search bar. The main content area displays search results for 'Sample_8xazama8'. The results list three entries, each with a title, description, and file information. A circular chart on the right shows the distribution of data. The 'User Information' section lists the user's name, contact, and principal investigator.

Synergist
Catalog | Explore | Collaborate

Home | Help | Contact | Login | Export | Dashboard | 7 items cart | Warn

Search Results - Found 3 matches

Quick Filter

Sample_8xazama8 | Title: Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course
Genotype:Wild-type | Organism:Human | Experiment:Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course | RNA type: Poly-A
RNA | Sequencing Type: Paired-End
Lab: Mello Lab | Conducted by: Sample User 8 | Facility: RTI
Data File Name: Project_SAA_8392_Custom.tar

Sample_8xazama8 | Title: Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course
Genotype:Wild-type | Organism:Human | Experiment:Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course | RNA type: Poly-A
RNA | Sequencing Type: Paired-End
Lab: Mello Lab | Conducted by: Sample User 8 | Facility: RTI
Data File Name: Project_SAA_8392_Custom.tar

Sample_8xazama8 | Title: Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course
Genotype:Wild-type | Organism:Human | Experiment:Sample Data - MicroRNA-133a overexpression effect on prostate cancer cell line time course | RNA type: Poly-A
RNA | Sequencing Type: Paired-End
Lab: Mello Lab | Conducted by: Sample User 8 | Facility: RTI
Data File Name: Project_SAA_8392_Custom.tar

Total number of records: 11/15 Data Sample of data belong to this subset

Details

User Information:

- Name: Sample User 8
- Contact: sample.user.8@unimed.edu
- Principal Investigator: Craig Mello
- Contact: craig.mello@unimed.edu



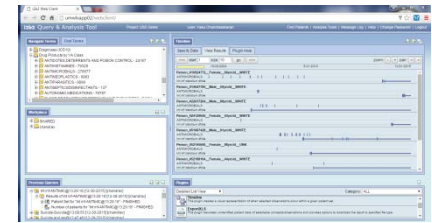
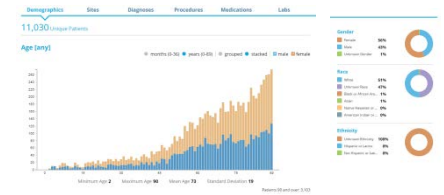
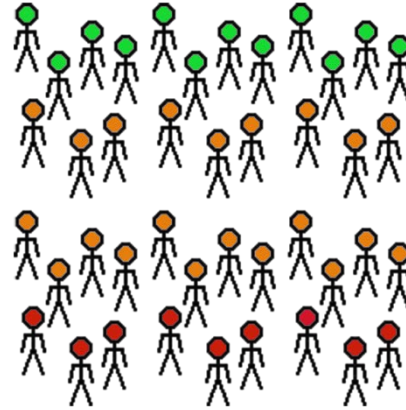
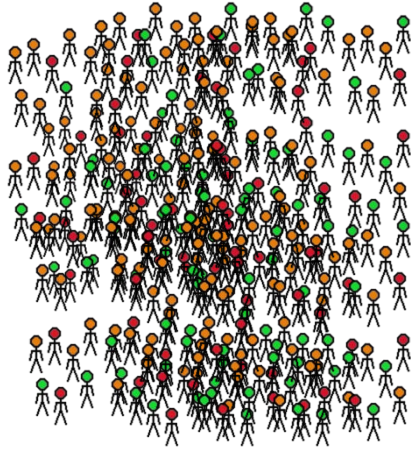
Note from the IRB

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, 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)

Integrative Data Ecosystem

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

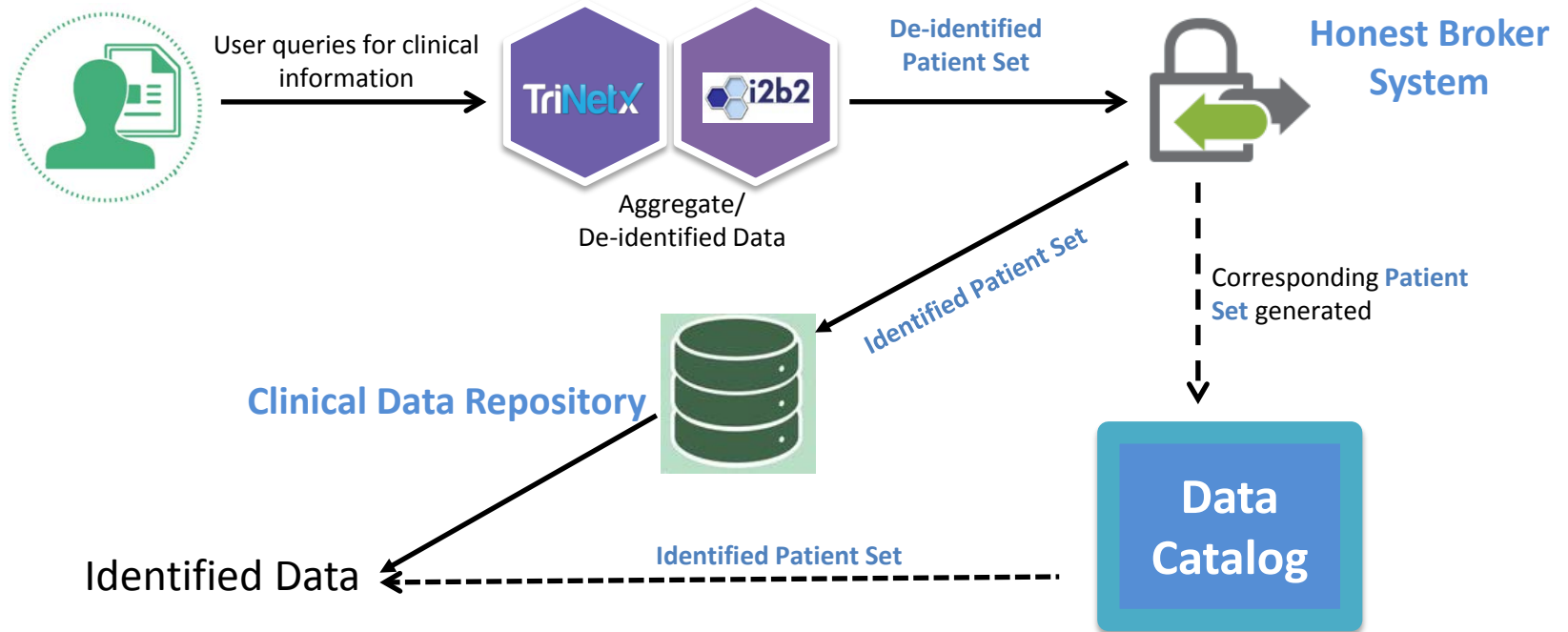


- Match patients to clinical trials
- Enable better and efficient design of clinical trials
- Identify meaningful points for intervention for 'continuous monitoring' scenarios
- Predict outcomes & Suggest treatment options



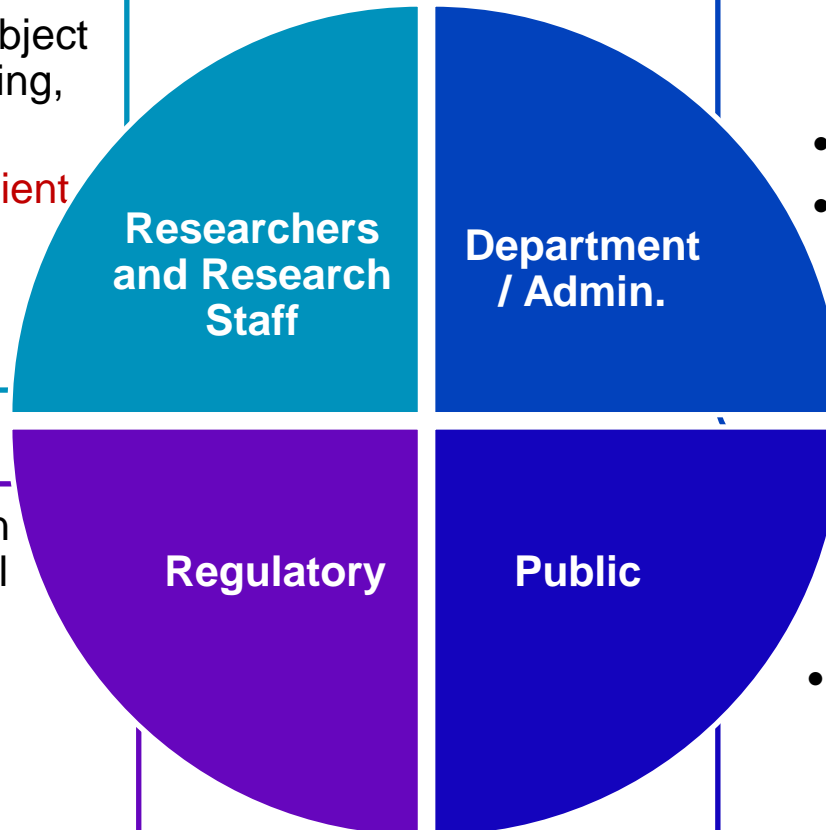
Translation
Enablers

Connect Research & Clinical Data



Oncore CTMS

- Improved communication across team and sites
- Automation support (subject calendars & re-consenting, sponsor invoicing)
- **EMR integration for Patient Safety**
- Grants reporting

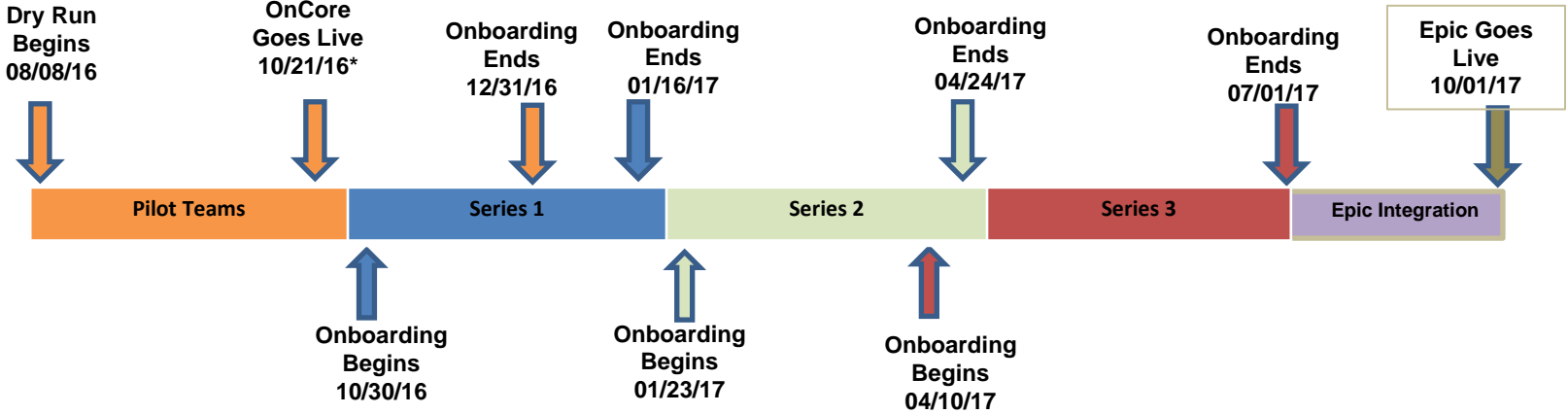


- Visibility to clinical research portfolio at different levels within the enterprise
- Resource allocation
- Effectively manage staff workload
 - Streamline budget negotiations
 - Billing compliance

- Standardized collection and reporting of clinical trial information
- Consistent tracking of safety data and staff credentials
- Email alerts and reminders

- Provides a searchable view of accurate and up-to-date open study information
- Enhance safety through EMR integration

Timeline for OnCore



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

* 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

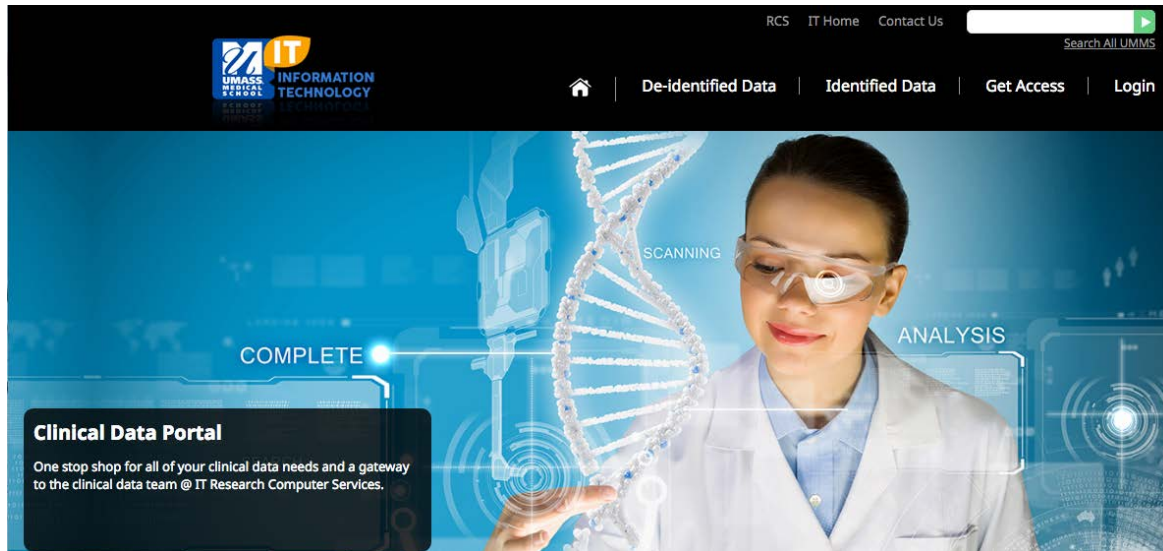
Beginning July 1, 2017, entry into and use of the CTMS will be a prerequisite for study initiation, account set-up and for ongoing financial management of the study.

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)*
- *Clinical Research Task Force (P. Muldoon & M. Koziel; T. Day, K Garabedian, D. Harlan, L. Harris, M. Hudlin, C. Jewell, M. Johnson, J. Mathew, J. Randolph, M. Rosen, T. Tarnowski, S. Tosi)*
- *Clinical Trial Management Steering Group (M. Koziel & J Mathew, J. Kay, J. Frazier, D. Habers. C. Keuker, P. Goden, D. Vazquez, G. Szabo, M. Johnson, V. Lee)*
- *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