APOLLO *a High Profile Use Case with Unique Challenges for the Cancer Research Data Commons*

15 May 2018

Kathleen M Darcy, PhD

Outline

- DOD Cancer Programs
- APOLLO Overview
- APOLLO Details
- APOLLO Data Flow, Standards, Integration and Analysis
- APOLLO Data Submission Options and Challenges

DOD Cancer Programs

NIH NATIONAL CANCER INSTITUTE

- Extramural Programs
- Intramural Programs

DEFENSE HEALTH AGENCY · DEPARTMENT OF DEFENSE · PRO CURA MILITIS

- **Extramural Programs**
  Congressionally Directed Medical Research Program
  - Breast Cancer Research Program
  - Ovarian Cancer Research Program
  - Prostate Cancer Research Program
  - Lung Cancer Research Program
  - Peer Reviewed Cancer Research Program
- **Intramural Programs**
  Uniformed Services University
  Murtha Cancer Center Centers of Excellence
  - USU Breast Cancer Center of Excellence
  - USU Gynecologic Cancer Center of Excellence
  - USU Prostate Cancer Center of Excellence

Provided by Chad Hamilton and Craig Shriver

Federal Precision Oncology Initiative of the National Cancer Moonshot

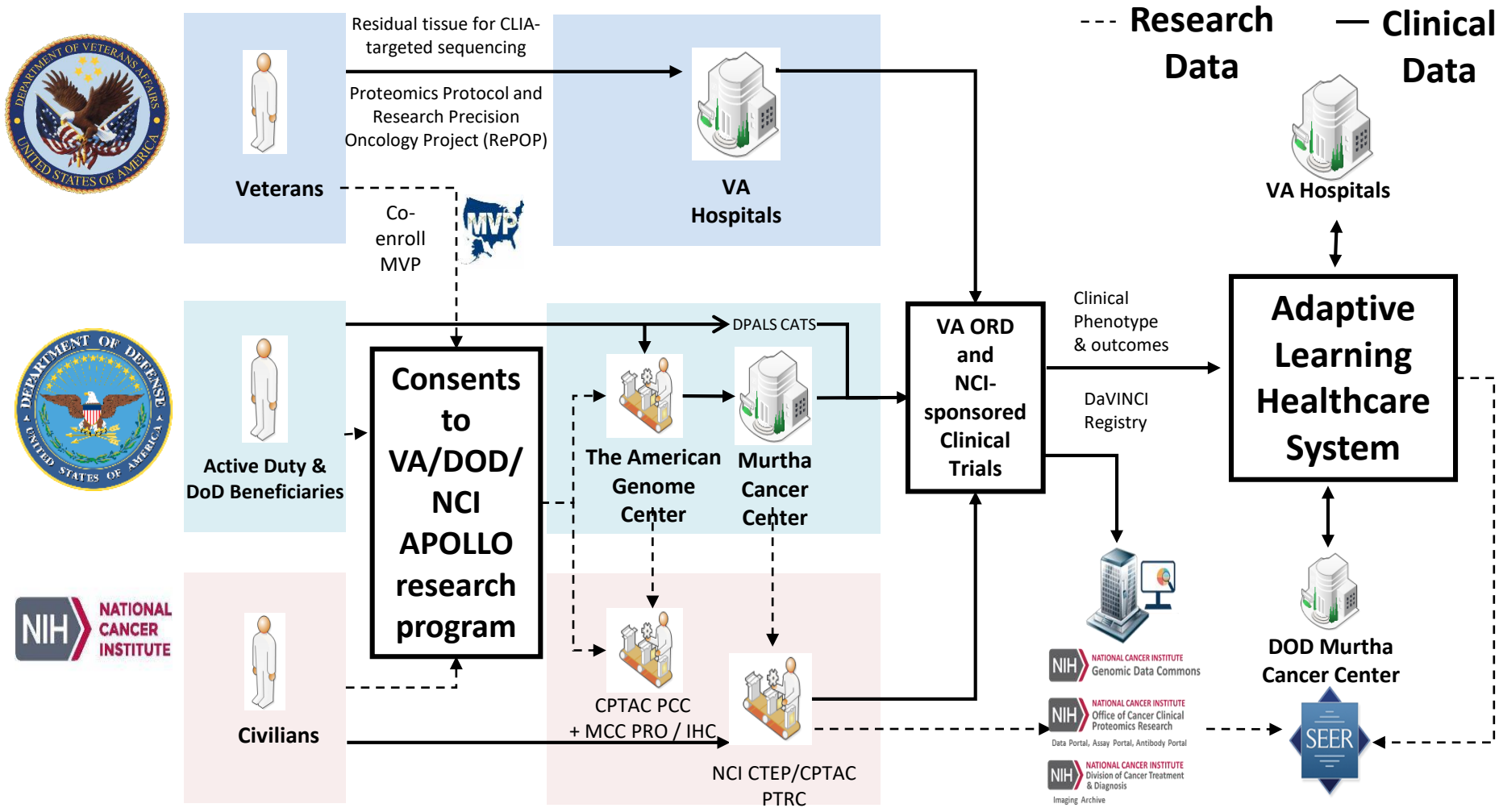
- **APOLLO**: **A**pplied **P**roteogenomics **O**rganizational **L**earning and **O**utcomes Consortium
  - APOLLO-1 Lung Cancer (existing cohort)
  - APOLLO-2 GYN Cancer (existing cohort)
  - APOLLO-3 Prostate Cancer (existing cohort)
  - APOLLO-4 Breast Cancer (existing cohort)
  - APOLLO-5 Pan Cancers (prospective cohort)
  - APOLLO-X To Be Determined

APOLLO Vision

A Federal Alliance between DOD, VA and NCI through strong research collaborations and partnerships that

- Optimizes federal resources
- Enhances cancer research and discoveries
- Reduces duplication
- Leverages technologies
- Enhances intellectual capital
- Increases education and training opportunities
- Uses advanced methods in proteogenomics to characterize and compare tumors
- Develops a deeper understanding of cancer biology
- Identifies potential therapeutic targets and pathways for cancer prevention, detection and intervention

*A*pplied *P*roteogenomics *O*rganizationa*L* *L*earning and *O*utcomes (**APOLLO**) Consortium



APOLLO Goals

- Develop and deploy predictive markers
- Advance target discovery and drug development
- Execute better matched clinical trials and trial designs
- Support an adaptive learning health system
- Implement evidence-based solutions for informed decision making, early cancer diagnosis, treatment and recovery
- Improve health and health care for active military, beneficiaries, veterans and civilians
  - Reducing cancer health disparities and costs
  - Improving readiness, patient experience, outcomes and survival

APOLLO Approach

Working Groups Aug 2016 to Sept 2017

- IRB Protocol Working Group
- Samples Working Group
- Data Working Group
- Technology Working Group

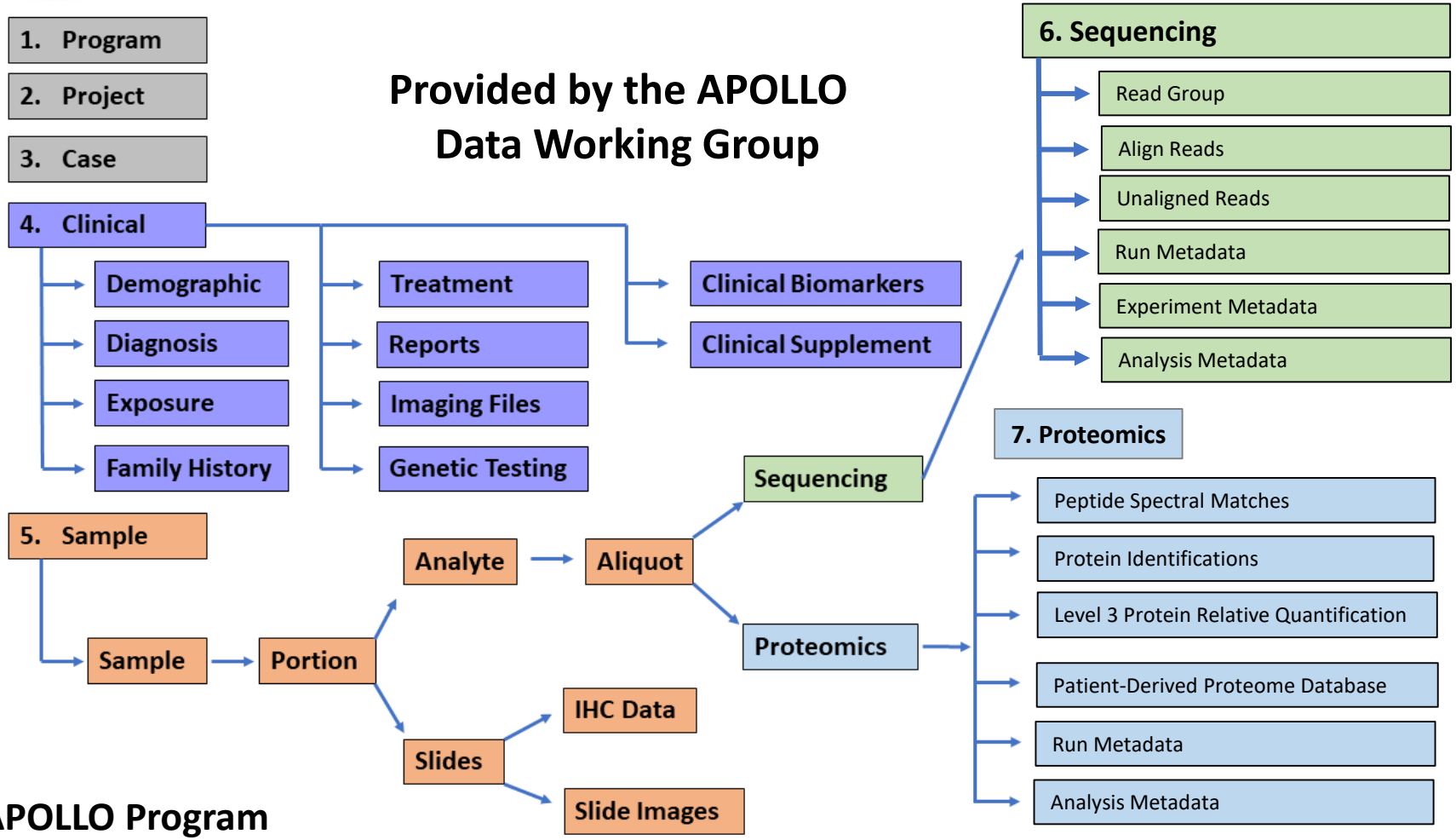
Working Groups Sept 2017 to Present

- Tissue Workflow Group
- Clinical Working Group
- Data Analysis Working Group
- Data Repository Working Group

Task Forces March 2018 to Present

- Publications
- Informatics Infrastructure

Incorporating APOLLO into the GDC Data Model



Provided by the APOLLO Data Working Group

APOLLO Program

- APOLLO Project (APOLLO-1, -2, -3, -4, -5, -X)
  - APOLLO Participants → APOLLO Clinical Data
    - APOLLO Aliquot IDs → Sample and either Sequencing or Proteomics Data

APOLLO Identifiers

- APOLLO Identifiers generated, distributed and managed by the Chan Soon-Shiong Institute of Molecular Medicine (CSSIMM) at Windber and registered with dbGAP
  - Participant Identifiers
    - Participant ID: AP-B3X7
    - 128-byte Global Unique Participant ID
  - Aliquot (Sample) Identifiers
    - Aliquot ID: AP-B3X7-KW
    - 128-byte Global Unique Aliquot ID

1M PARTICIPANT IDS AND 1K ALIQUOT IDS AVAILABLE FOR APOLLO.

Provided by Hai Hu and Leonid Kvecher

APOLLO Workflow Overview

Retrospective and Prospective Protocols

- Acquire APOLLO participant and aliquot IDs
- Review consent or recruit and consent patients
- Acquire specimens, clinical and patient-reported data linked to APOLLO IDs
- Perform QA, modality and domain-specific reviews, resolve queries, select cases for testing, recode data
- Process specimens, prepare analytes and generate proteogenomic data
- Submit clinical and tissue imaging to and recover feature annotation from the Cancer Imaging Archive
- Aggregate and analyze level 3 data with clinical and patient-reported data using the APOLLO Data Warehouse and NCI Jamboree site
- Share data with NCI Cancer Research Data Commons

APOLLO Protocols and Priorities

APOLLO 1-4

100-300 Cases by Site

- **APOLLO-1 Lung Cancer** *from the DOD CDMRP LCBRN and the VA*
- **APOLLO-2 GYN Cancer** *from the DOD GYN Cancer COE*
- **APOLLO-3 Prostate Cancer** *from the DOD Prostate Cancer COE*
- **APOLLO-4 Breast Cancer** *from the DOD Breast Cancer COE*

APOLLO-5 Pan Cancers

Estimated Cases/year by site

- GYN: 300-400
- Breast: 150-200
- Prostate: 50-100
- Colon/GI: 50-100
- ENT/Thyroid: 50-100
- Kidney: 25-50
- Lung: 25-50
- Brain: 10-20
- Sarcomas: 10-20
- Lymphoid: 10-20

from MCC Sites and COEs in GYN, Prostate and Breast Cancer

Projected accrual estimated to be ~8,000 cases

Priorities:

1. **Active Duty**
2. **Minorities**
3. **High priority cancers and cohorts**
   - **Aggressive or rare subtypes**
   - **Metastatic disease**
   - **Recurrent or persistent disease**
   - **Resistant phenotype**

APOLLO Sites

- **CDMRP Lung Cancer Biospecimen Resource Network** (LCBRN) for **APOLLO-1**
  - Civilian Sites: University of Virginia (UVA), Medical University of South Carolina, and Washington University of St Louis
- **VA** contributes existing Lung Cancer Cases for **APOLLO-1**
- **Gynecologic Cancer Center of Excellence** Tissue and Data Acquisition Network (TDAN) for **APOLLO-2** and **APOLLO-5**
  - Civilian Sites: Inova, Duke, OSU, Roswell Park and UVA
- **Prostate Cancer Center of Excellence** for **APOLLO-3** and **APOLLO-5**:
  - DOD Site: WRNMMC
- **Breast Cancer Center of Excellence** for **APOLLO-4** and **APOLLO-5**
  - DOD Site: WRNMMC
  - Civilian Sites: Joyce Murtha Breast Care Center and Anne Arundel Medical Center
- **Murtha Cancer Center (MCC)** Biobank for **APOLLO-5**
  - DOD Sites: WRNMMC, Ft. Bragg, Portsmouth, Keesler, San Diego, Madigan, Fort Belvoir, San Antonio, William Beaumont El Paso
  - VA Site: VA Palo Alto
  - Civilian Site: Anne Arundel Medical Center
- Additional VA, DOD and Civilian Sites will be considered for **APOLLO-5** and the **APOLLO-X series**

APOLLO Facilities

- **CAP Accredited Biorepositories**
  - Murtha Cancer Center (MCC) Biobank
  - Chan Soon-Shiong Institute of Molecular Medicine (CSSIMM) at Windber
  - Women's Health Integrated Research Center (WHIRC) at Inova
  - Center for Prostate Disease Research (CPDR)
- **Processing Centers**
  - CSSIMM at Windber
  - Pathology Research Center at the WHIRC at Inova
- **Analytical Facilities**
  - The American Genome Center (TAGC) at USUHS
  - The Murtha Cancer Center Clinical Proteomics Platform
  - The Clinical Proteomics Tumor Analysis Consortium (CPTAC)
  - The Murtha Cancer Center Immunohistochemistry (IHC) Laboratory

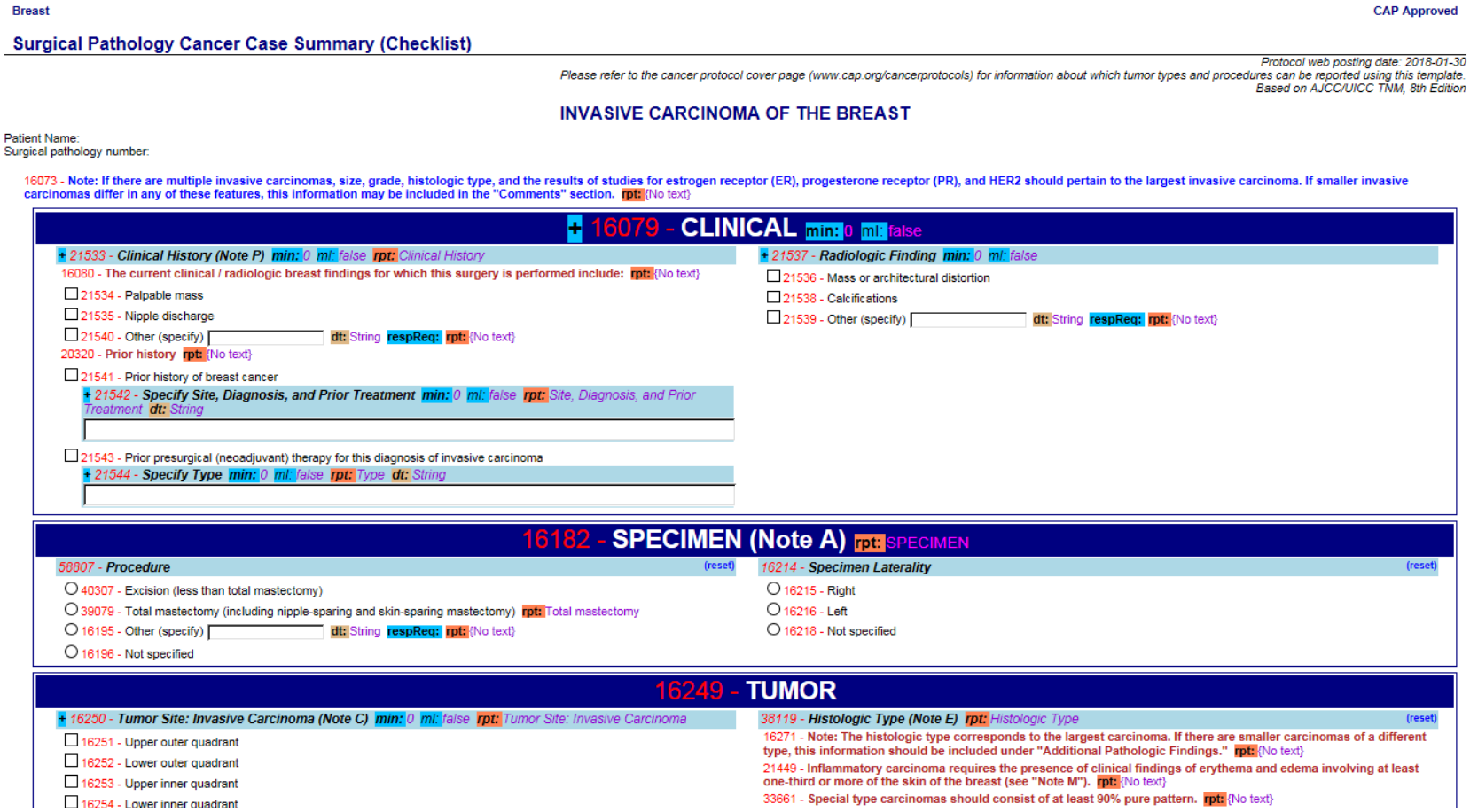
APOLLO Data Model

1. PROGRAM | 2. PROJECT | 3. CASES | 4. CLINCIAL | 5. SAMPLE | 6. SEQUENCING | 7. PROTEOMICS

- **4. Clinical Data**
  - **NCI Clinical Data Elements (GDC)**
    - Demographic
    - Diagnosis
    - Exposure
    - Family History
    - Treatment
  - **APOLLO Domain-Specific Elements**
    - CAP eCC Pathology Checklist
    - Detailed Treatment & Follow Up
    - CLIA Gene Testing & Biomarkers
    - Medical Reports & Imaging Files
    - Patient Reported Measures
- **5. Sample**
  - **NCI Sample Data Elements (GDC)**
  - **CAP eCC Sample Annotation by MCC and JPC Pathologists**
  - **Clinical Image Annotation**
    - Radiologist Assessed
    - Computer Generated
  - **Tissue Image Annotation**
    - Pathologist Assessed
    - Computer Generated
- **6. Sequencing**
  - Whole Genome Sequencing
    - Germline DNA
    - Tumor DNA
  - RNA Sequencing
  - Read Group
  - Align Reads
  - Unaligned Reads
  - Run Metadata
  - Experiment Metadata
  - Analysis Metadata
- **7. Proteomics**
  - TMT Global Proteomics
  - TMT Phospho Proteomics
  - Reverse Phase Protein Array
  - MRM Proteomics
  - Peptide Spectral Matches
  - Protein Identifications
  - Level 3 Protein Relative Quantification
  - Patient-Derived Proteome Database
  - Run Metadata
  - Analysis Metadata

APOLLO Pathology Data

- Synoptic pathology data and biomarkers captured using XML forms and codes provided by CAP
  - Accept XML exports from CoPath Plus imported into harmonized eCC XML forms



Breast

CAP Approved

Surgical Pathology Cancer Case Summary (Checklist)

Protocol web posting date: 2018-01-30
Please refer to the cancer protocol cover page (www.cap.org/cancerprotocols) for information about which tumor types and procedures can be reported using this template.
Based on AJCC/UICC TNM, 8th Edition

INVASIVE CARCINOMA OF THE BREAST

Patient Name:
Surgical pathology number:

16073 - **Note: If there are multiple invasive carcinomas, size, grade, histologic type, and the results of studies for estrogen receptor (ER), progesterone receptor (PR), and HER2 should pertain to the largest invasive carcinoma. If smaller invasive carcinomas differ in any of these features, this information may be included in the "Comments" section.** rpt: {No text}

+ 16079 - CLINICAL min: 0 ml: false

+ *21533 - Clinical History (Note P)* min: 0 ml: false rpt: Clinical History

16080 - **The current clinical / radiologic breast findings for which this surgery is performed include:** rpt: {No text}

- ☐ 21534 - Palpable mass
- ☐ 21535 - Nipple discharge
- ☐ 21540 - Other (specify) dt: String respReq: rpt: {No text}

20320 - **Prior history** rpt: {No text}

- ☐ 21541 - Prior history of breast cancer
  - + *21542 - Specify Site, Diagnosis, and Prior Treatment* min: 0 ml: false rpt: Site, Diagnosis, and Prior Treatment dt: String
- ☐ 21543 - Prior presurgical (neoadjuvant) therapy for this diagnosis of invasive carcinoma
  - + *21544 - Specify Type* min: 0 ml: false rpt: Type dt: String

+ *21537 - Radiologic Finding* min: 0 ml: false

- ☐ 21536 - Mass or architectural distortion
- ☐ 21538 - Calcifications
- ☐ 21539 - Other (specify) dt: String respReq: rpt: {No text}

16182 - SPECIMEN (Note A) rpt: SPECIMEN

58807 - Procedure (reset)

- ○ 40307 - Excision (less than total mastectomy)
- ○ 39079 - Total mastectomy (including nipple-sparing and skin-sparing mastectomy) rpt: Total mastectomy
- ○ 16195 - Other (specify) dt: String respReq: rpt: {No text}
- ○ 16196 - Not specified

16214 - Specimen Laterality (reset)

- ○ 16215 - Right
- ○ 16216 - Left
- ○ 16218 - Not specified

16249 - TUMOR

+ *16250 - Tumor Site: Invasive Carcinoma (Note C)* min: 0 ml: false rpt: Tumor Site: Invasive Carcinoma

- ☐ 16251 - Upper outer quadrant
- ☐ 16252 - Lower outer quadrant
- ☐ 16253 - Upper inner quadrant
- ☐ 16254 - Lower inner quadrant

38119 - Histologic Type (Note E) rpt: Histologic Type (reset)

16271 - **Note: The histologic type corresponds to the largest carcinoma. If there are smaller carcinomas of a different type, this information should be included under "Additional Pathologic Findings."** rpt: {No text}

21449 - **Inflammatory carcinoma requires the presence of clinical findings of erythema and edema involving at least one-third or more of the skin of the breast (see "Note M").** rpt: {No text}

33661 - **Special type carcinomas should consist of at least 90% pure pattern.** rpt: {No text}

Provided by Justin Wells and Leonid Kvecher

APOLLO CLIA Gene Panel Data

- Upload redacted CLIA Gene Panel Testing Reports in pdf into the Data Tracking System
- Parse and import CLIA Gene Panel Testing Report Findings in XML into the DTS

DOD Illumina True Sight Tumor Panel

- TST15: AKT1, BRAF, EGFR, ERB2, FOXL2, GNA11, GNAQ, KIT, KRAS, MET, NRAS, PDGFRA, PIK3CA, RET and TP53
- TST170: Assessment of DNA and RNA for fusions, splice variants, insertions/deletions and single-nucleotide variants (SNVs), and amplifications in one assay.

VA Gene Panel Testing Vendors

- Personalis ACE CancerPlus: 181 cancer genes
- PGDx CancerSelect 125 Test: 125 genes

CLIA testing drives treatment selection for FDA indications and participation in Clinical Trials

- Gene panel reports include variant annotations (PDF or XML format)
  - ✓ Level of evidence
    - ❑ **Tier I**: Variant with Strong Clinical Significance (Level A or Level B)
    - ❑ **Tier II**: Variant with Potential Clinical Significance (Level C or Level D)
    - ❑ **Tier III**: Variant with Unknown Significance
    - ❑ **Tier IV**: Benign or Likely Benign Variant
  - ✓ Summary of results
  - ✓ Molecular function
  - ✓ Incidence and role in disease
  - ✓ Effect on drug sensitive or resistance
  - ✓ Therapies targeting the variant
  - ✓ Trials prioritized by clinical specificity
  - ✓ Trials prioritized by region

Provided by Joel Moncur (DOD) and Brett Johnson (VA)

APOLLO Patient-Reported Data

- **CLASSIC EPIDEMIOLOGIC ASSESSMENTS**:
  1. **Patient Demographics** including race, ethnicity, sex, marital status, education, employment and military service
  2. **Medical History** regarding health conditions, any prior cancer diagnoses and treatments, as well as height and weight
  3. **Physical Activity** for 12 months prior to the current diagnosis
  4. **Alcohol History** in your entire life and currently
  5. **Tobacco Products** in your entire life and currently
  6. **Work Environment** including occupations, exposures and deployments
  7. **Family Cancer History** for blood relatives and ½ blood relatives
  8. **Reproductive History** for females

Completed during an interview with the Research Associate

Provided by Sarah Sakura and the APOLLO IRB Protocol Working Group

APOLLO Patient-Reported Data

- **PROMPT ASSESSMENTS**:
  1. **FACT-G Quality of Life:** physical, social/family, emotional and functional well-being
  2. **Global Health** evaluated using the PROMIS Global Health 10 v1.2
  3. **Pain and Fatigue** using the PROMIS PAIN 3a and PROMIS Fatigue 4a
  4. **Stress, Anxiety and Depression** using the NIH Toolbox Perceived Stress 10 instrument, PROMIS Anxiety 4a and PROMIS Depression 4a instruments
  5. **Symptoms** evaluated using the FACT NTX-4 for neurotoxicity, PROMIS Cognitive Function 4a for chemobrain and PROMIS Sleep Disturbance 4a instruments
  6. **Support for Daily Living** using the PROMIS Instrumental Support v2.0 instrument

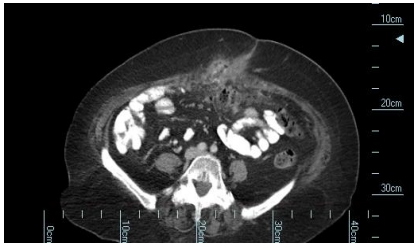
- **FOCUS ASSESSMENTS**:
  1. **FACT Cancer Specific Concerns** from FACIT.org
  2. **Barriers to Care** using a customized instrument
  3. **Patient Preferences** using a customized instrument
  4. **Events** using the Impact of Events (IES) instrument (PTSD assessment tool)
  5. **Financial Well-Being** using the FACT Financial Toxicity instrument
  6. **Spiritual Well-Being** using the FACT Spiritual Well-Being instrument

Provided by the APOLLO PROs Subgroup

APOLLO Clinical Imaging

Textural Feature Annotation Team to be led by Evis Sala MD, PhD

Pre-NACT



Post-NACT pre-op

8 days s/p IDS

First Recurrence

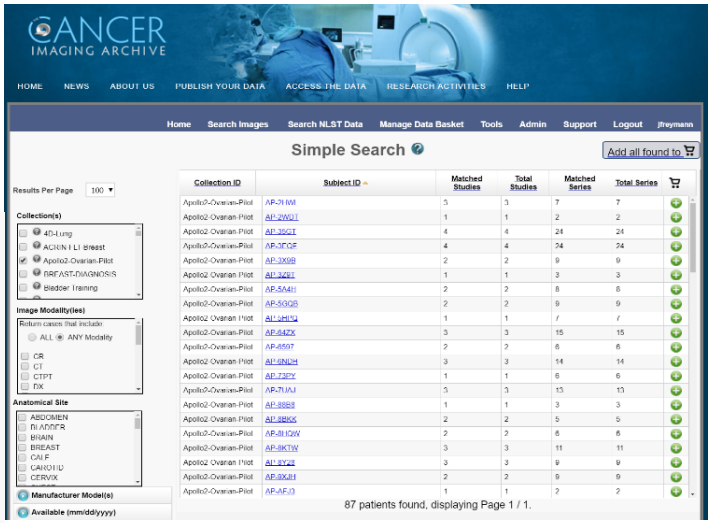
Full Cohort	
Inova	30
Duke	29
OSU	44

Type of Scans	Count
CT	232
PET	2
MRI	5
US	2
X-Ray	9

Time Points	Cases
Pre-Op	55
Pre-NACT	16
Post-NACT	13
Post-Op	48
Post-ACT	31
Recurrence 1	50
Multiple Recurrences	12
Last image available	14

Cases Curated	90
Cases Published	87

Total Scans	250
Total Files	154,672

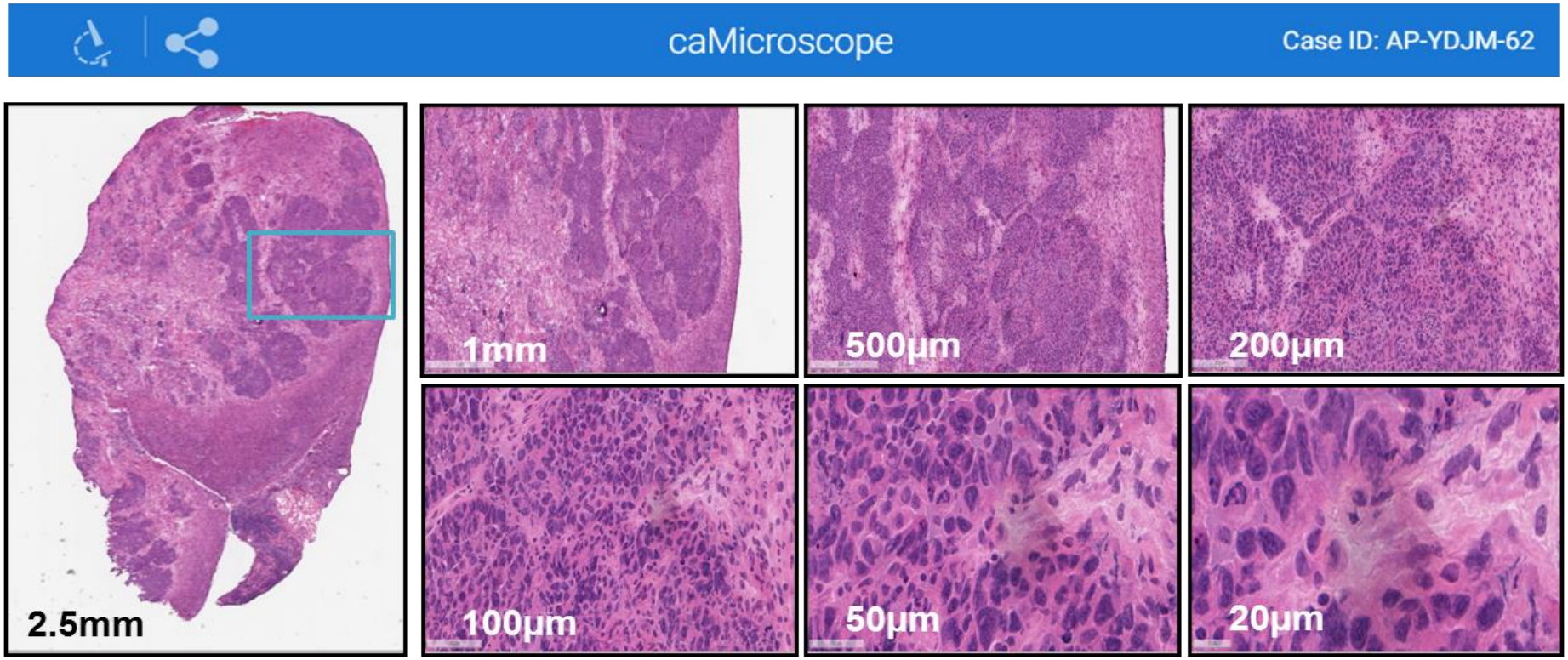


Study Instance UID	Description	Date	Add This Study To Basket
1.3.6.1.4.1.14519.5.2.1.5472.5801.313540898271840521570979738896	CT Abdomen and Pelvis W/Contrast=A	Baseline	

Series	Description	Modality	Manufacturer	Images	Thumbnails	Cine mode	DICOM
....5808503176	Topogram 0.6 T20s	CT	SIEMENS	1		NA	
....1757565394	Abdomen 5.0 I40f 2	CT	SIEMENS	101			
....7297979203	Coronal	CT	SIEMENS	124			

Provided by the GYN-COE and TCIA APOLLO-2 TEAM

APOLLO Tissue Imaging



- GYN-COE has uploaded reference hematoxylin and eosin (H&E) - stained frozen tissue sections for APOLLO2 high grade serous ovarian cancers (HGSOC, n=105) to TCIA.

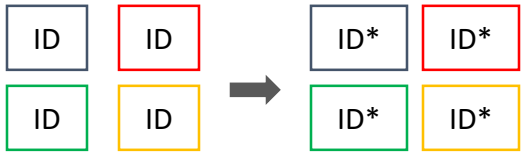
- Representative APOLLO2 case in TCIA portal extracted using caMicroscope software (Ashish Sharma, Emory University).

Provided by the GYN-COE and TCIA APOLLO-2 TEAM

VA Imaging SOP

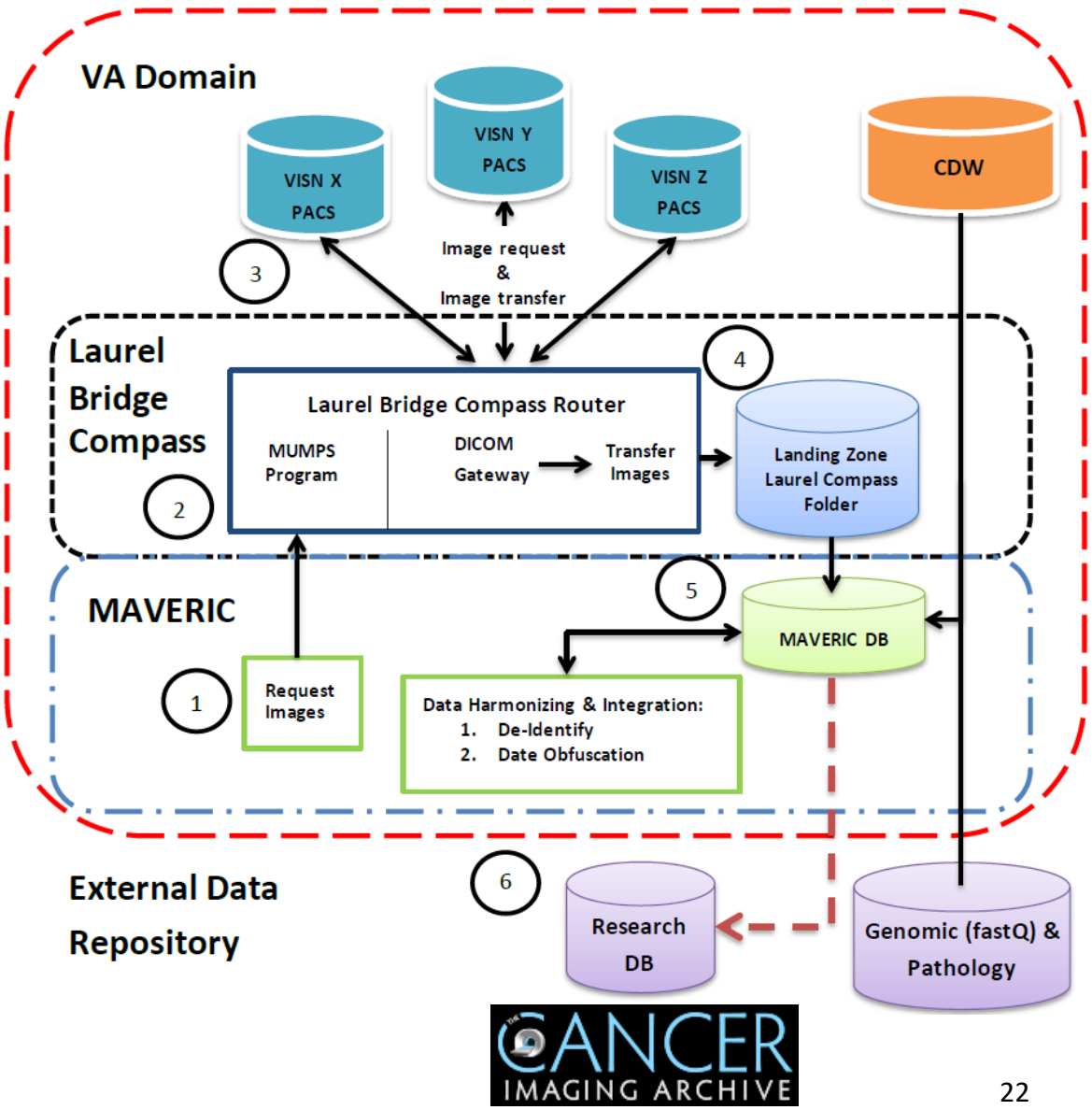
Workflow

1. List of Images is compiled
2. List uploaded to LBC
3. LBC Router forage across VA (VISNs) for images
4. Located images are pulled to LBC Router and pushed to LZ (LBC Folder)
5. Data is harmonized & integrated via PO approved process:



Clinical Data is Nationwide De-Id
Images De-Id via CTP & POSDA
Genomic & Pathology De-Id

6. Encrypted data is sent or uploaded to External VA repository i.e., TCIA -The Cancer Imaging Archive (NCI).



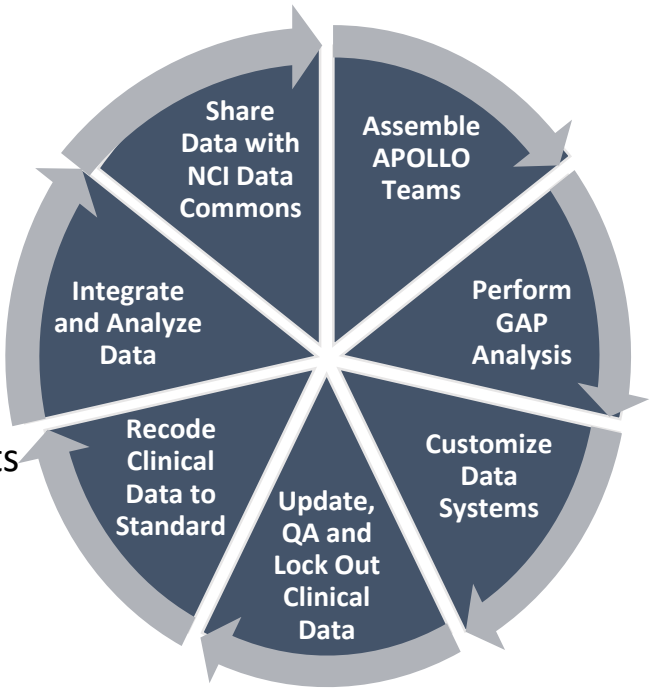
Luis E. Selva, Ph.D., MS, BSME, Associate Director of Imaging Informatics, MAVERIC, VA Boston Healthcare System

APOLLO Data Standards

Data Tracking System (DTS) for APOLLO

A user-friendly agile system with role-based access, logging and audit reporting, Smart forms and automated processes

- ❑ Register cases and IDs
- ❑ Consent data & restrictions
- ❑ Surgical data
- ❑ Follow up and outcomes
- ❑ Epidemiology Data
- ❑ Patient Reported Outcomes
- ❑ Imaging Features
- ❑ pdf Reports
- ❑ Central Path Review and annotation of tumor tissues
- ❑ Parse and accept XML imports
  - ➢ CAP eCC for 25+ cancers



Utilize Harmonized Processes and Standardized Documentation

- ❑ Develop scripts for standard recoding of clinical data from LCBRN, GYN-COE, CPDR and CBCP
- ❑ Utilize semantic annotation tools to add ontology and harmonize to new standards
- ❑ QC, reconcile, finalize and lock out data
- ❑ Continue to follow patients to update disease status, cancer Tx, CLIA testing, vital status, codth

Follow Best Practices

- ❑ Create dbGAP account
- ❑ Coordinate with the NCI points of contact
- ❑ Upload data into the NCI CRDC and TCIA

Provided by the APOLLO Clinical Working Group

APOLLO Workflow and Submission

Retrospective and Prospective Protocols

Source Sites

- MCC Biobank Sites DOD, VA
- Prostate Cancer COE Site DOD
- Breast Cancer COE Sites DOD, Civilian
- GYN Cancer COE Sites Civilian
- Lung Cancer Sites Civilian, VA

KEY IDENTIFIERS

Subject ID: AP-B3X7 with a 128-byte Global Unique Subject ID

Aliquot ID: AP-B3X7-KW with a 128-byte Global Unique Aliquot ID

APOLLO Repositories and Processing Centers

CSSI at Windber
FW, DTS & DW4TR

APOLLO IDs and Samples →

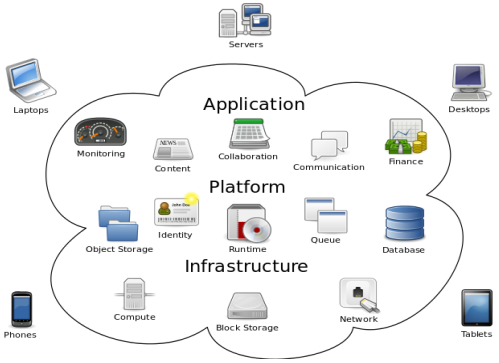
← APOLLO Clinical and Sample Data

GYN-COE at WHIRC
FW & REDCap

Analytic Facilities

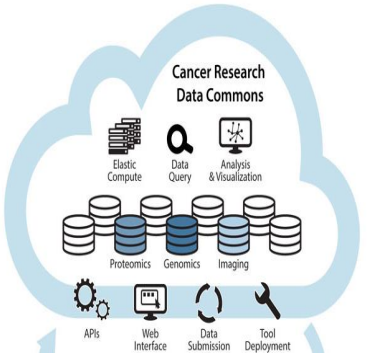
- The American Genome Center (TAGC) at USUHS
- MCC IHC Laboratory
- MCC Clinical Proteomics Platform
- Clinical Proteomic Tumor Analysis (CPTAC)

APOLLO Integrated Analyses



APOLLO Data Warehouse and the NCI Jamboree Site

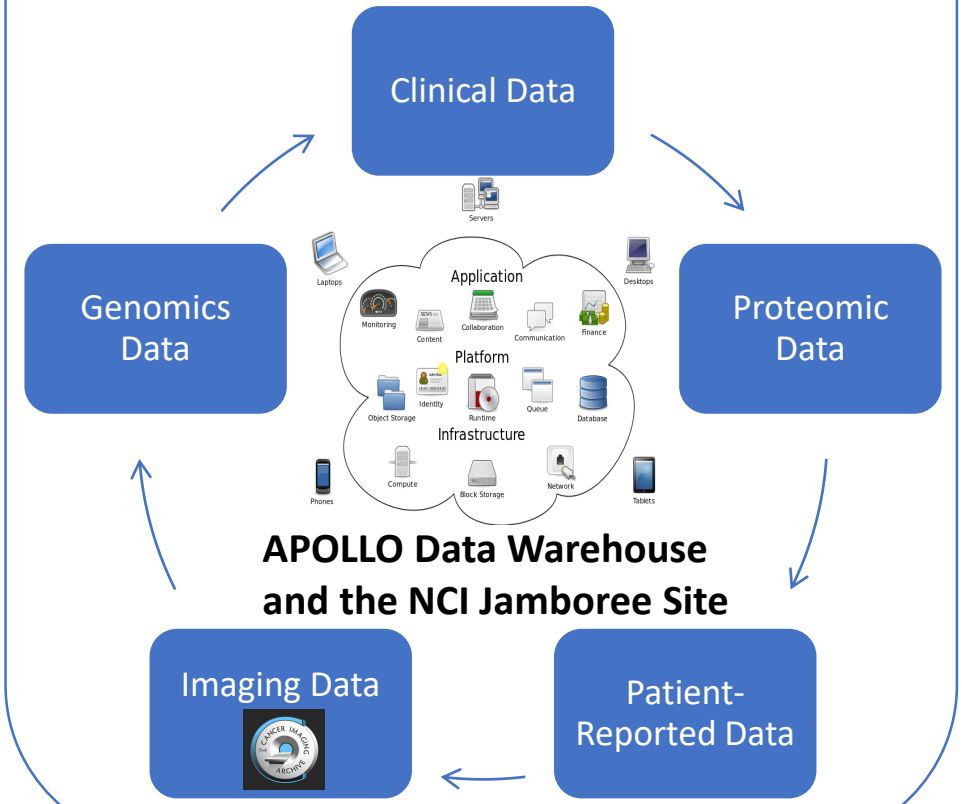
NCI Cancer Research Data Commons



APOLLO DATA SHARING

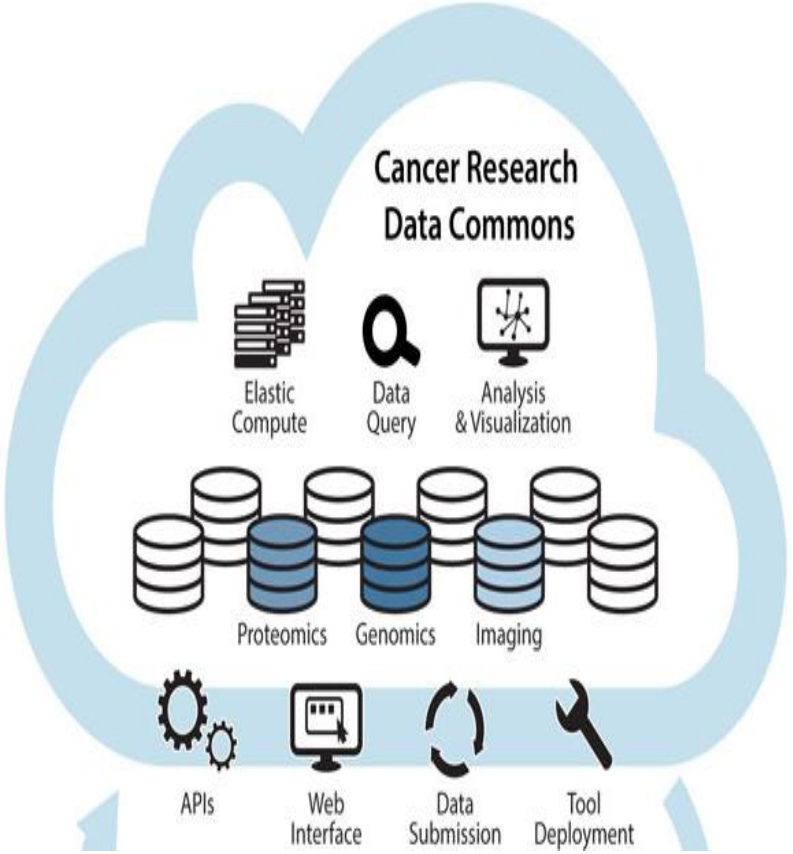
APOLLO Research Network

Data Collection, Generation, Harmonization, Integration, Analysis, and Publication



APOLLO Data Warehouse and the NCI Jamboree Site

NCI Cancer Research Data Commons (CRDC)



Cancer Data Aggregator (CDA)

During Data Submission from the APOLLO Research Network to the NCI CRDC

- **Will APOLLO data need to be partitioned, recoded and reaggregated?**
- **Who will be responsible: submitter and/or the data commons?**

Acknowledgements

- Jerry Lee
- Jennifer Lee
- Christopher Moskaluk
- Warren Kibbe
- Lynn Penberthy
- Sean Hanlon
- Mickey Williams
- Neil Spector
- Henry Rodriguez
- Amanda Paulovich
- Frank Meng
- Lou Fiore
- Luis Selva
- Danne Elbers
- Brett Johnson
- John Freymann
- Paula Jacobs
- Denise Warzel
- Justin Kirby
- Craig Shriver
- Hai Hu
- Stella Somiari
- Leonid Kvecher
- Justin Wells
- Joel Moncur
- Clesson Turner
- Mary Lou Cutler
- Joseph Vockley
- Matthew Wilkerson
- Clifton Dalgard
- Harvey Pollard
- Terry Rauch
- Fiona Renalds
- Sarah Sakura
- Jamie Bonne
- Izumi Hinkson
- Daoud Meerzaman
- Tanja Davidson
- G. Larry Maxwell
- Thomas Conrads
- Nicholas Bateman
- Chad Hamilton
- Kerri Cronin
- James Bates
- Isabell Sesterhenn
- Shiv Srivastava
- Inger Rosner
- Jennifer Cullen
- Stanley Lipkowitz
- Denise Wright
- Jeffrey Hooke
- Al Kovatich
- Jeremy Perkins
- Lari Wenzel
- Frankie Cozzens-Phillips
- Kelli Ruiz
- Charles Goldthwaite

Special thanks to Leonid Kvecher, Hai Hu, Justin Wells, John Freymann, Justin Kirby, Joel Moncur, Brett Johnson, Nick Bateman, G. Larry Maxwell, Tom Conrads, and Craig Shriver for contributions to this presentation