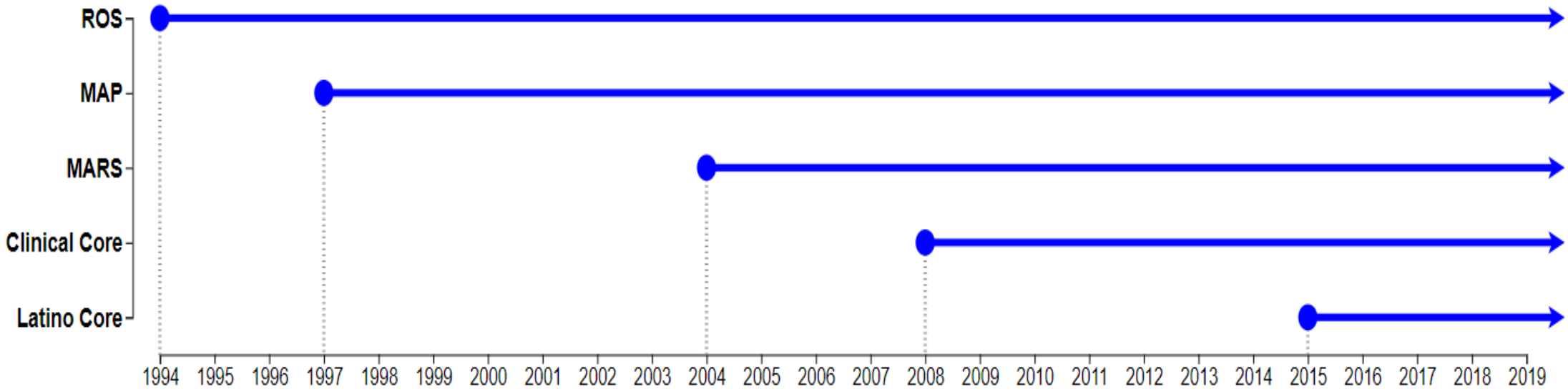
Spring ADC Meeting 2019

Data and Specimen Sharing
Tracking and Documentation

John Gibbons, MS
Data Manager / Manager Information System
Rush Alzheimer’s Disease Center

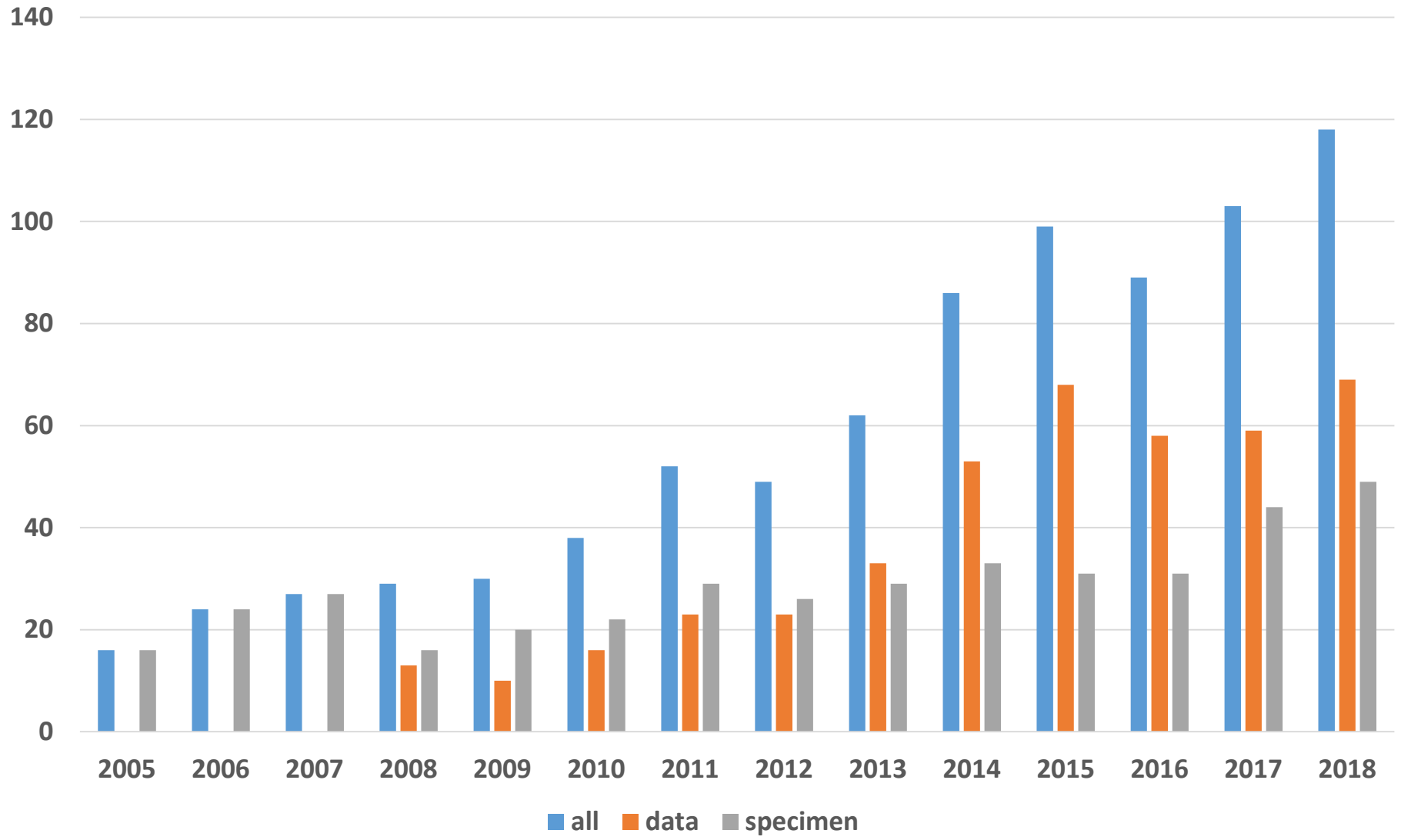
RADC Cohort Studies



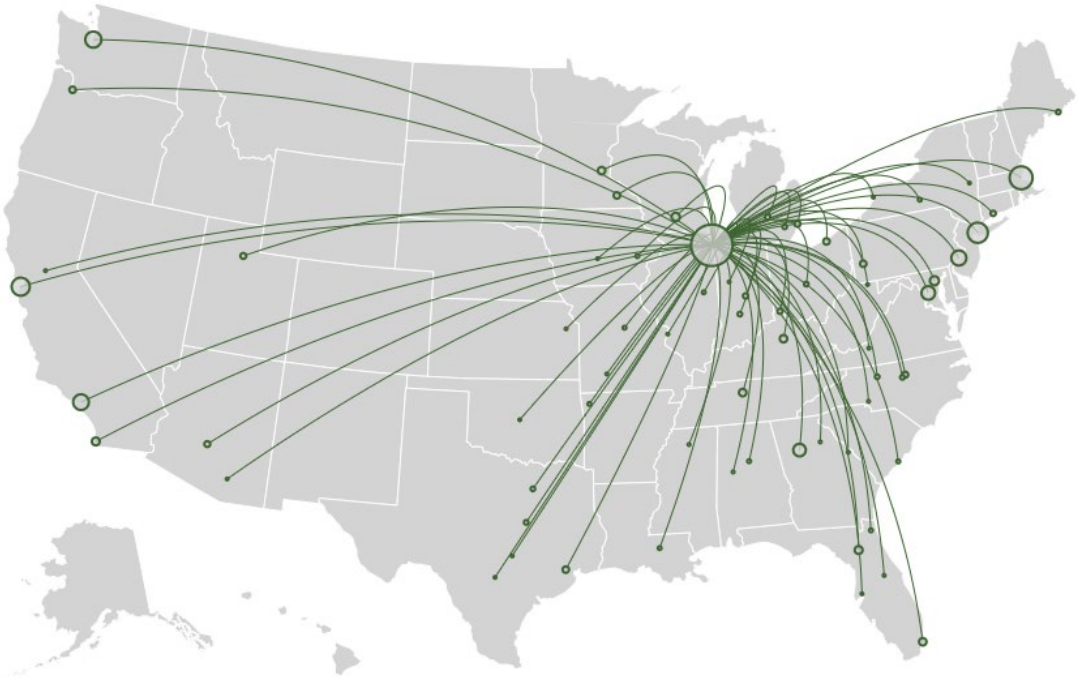
Data & Specimens

Study	Enrollment	Autopsy
ROS	1460	777
MAP	2159	828
MARS	758	27
Clinical Core	373	67
Latino Core	221	1
Total	4,971	1,700

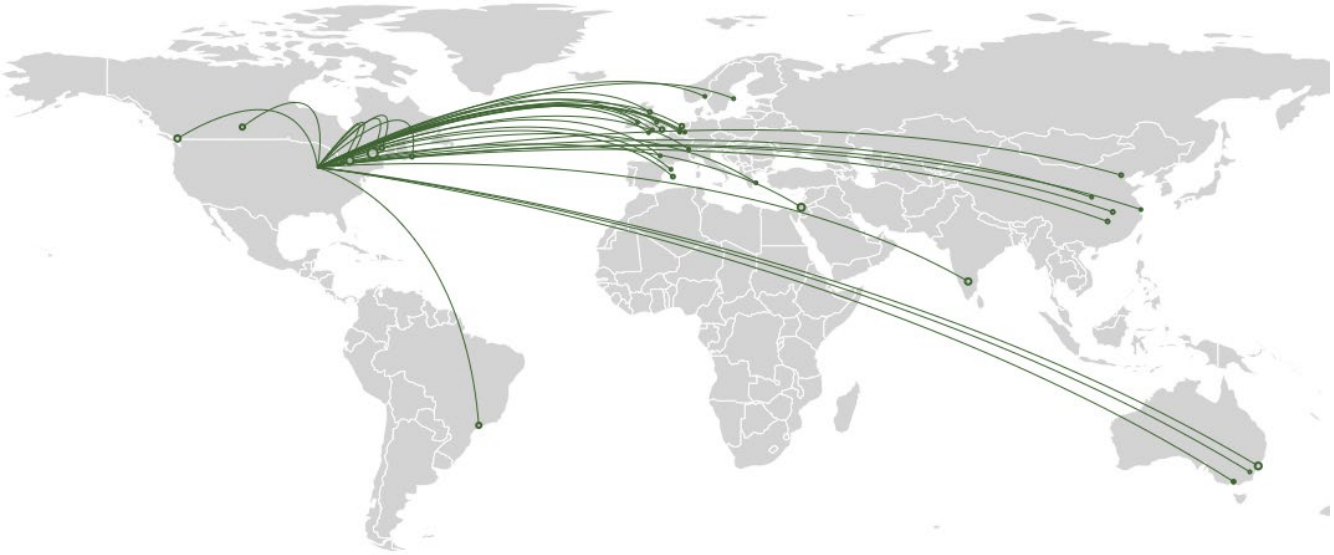
Data and Specimen Requests by year



National



International

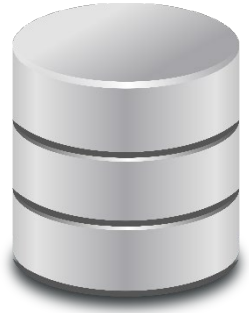


- Resource Sharing Hub
- Request Review Process
- DUA & MTA Tracking
- Specimen Selection

- Specimen Distributions
- Plating of Aliquots
- Generation of Data Sets and Codebooks

- Project States
- Tracking Shipments
- Tracking Specimen Distribution

Central Relational Database – SQL



- Participant information
  - Unique identifier
  - Consent to studies and repositories
- Research data
  - Cognitive, interview, device, MRI, -omics
- Computed variables
  - Documented
  - Processed from item-level data
  - Evolve over time
- External Requests for data, specimens
- Biospecimen inventory
- Aggregated data for analytics

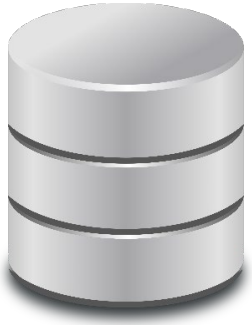
Central Relational Database – SQL

Intake of Outside Requests

RADC Research Resource Sharing Hub

Laboratory Processes and Specimen Shipping

Freezerworks



Compute Variables / Aggregate Data

sas SQL

Internal Website

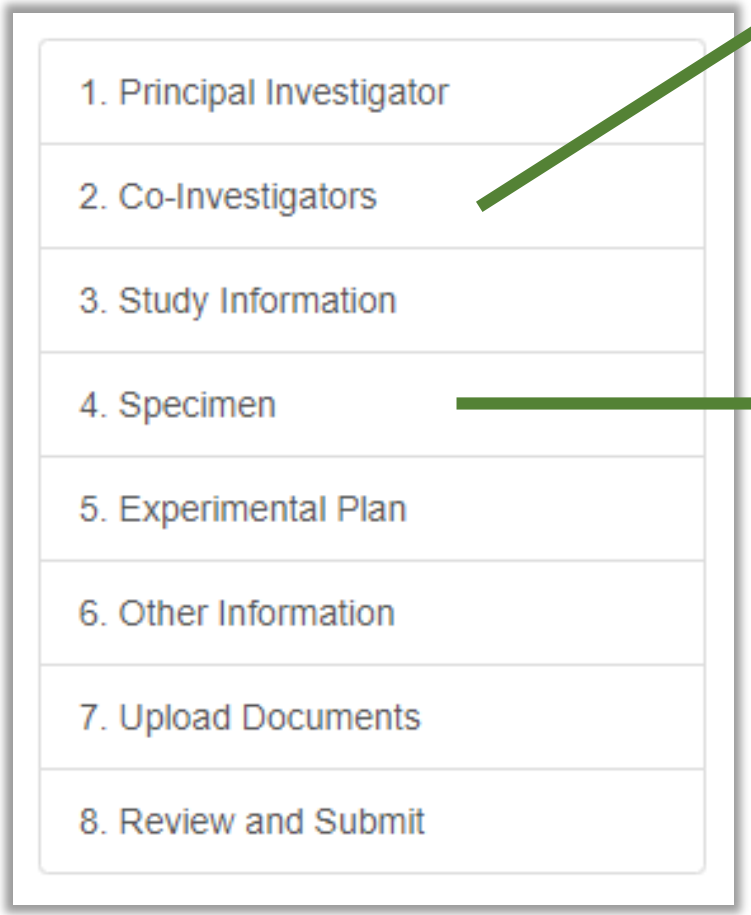
Study Data Collection

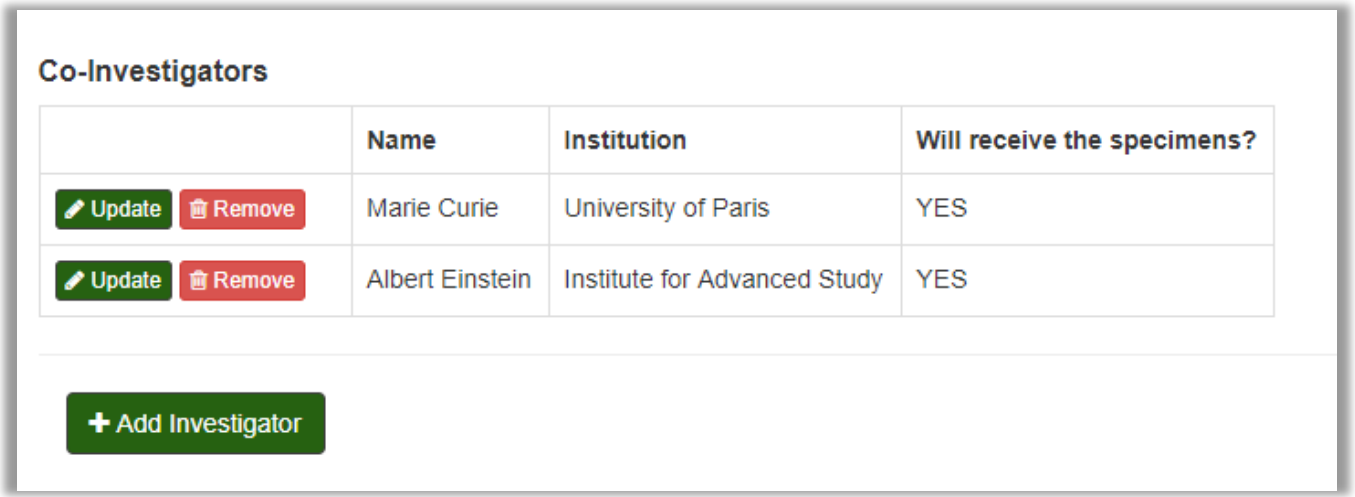
- Laptops – computer assisted interviewing forms
- Web-based forms

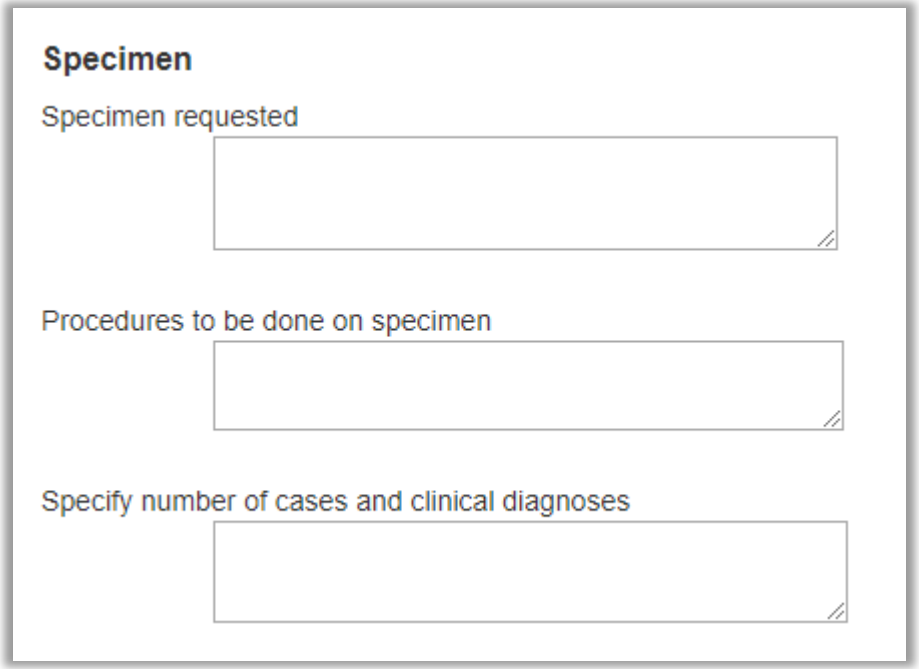
Research Resource Sharing Hub - www.radc.rush.edu

- Documentations
  - Studies / Sub-studies
  - Variables
  - -omics
  - Imaging
- Data Availability Reports
- Requests
  - Biospecimens
  - Limited Data Sets
  - Subject Recruitment

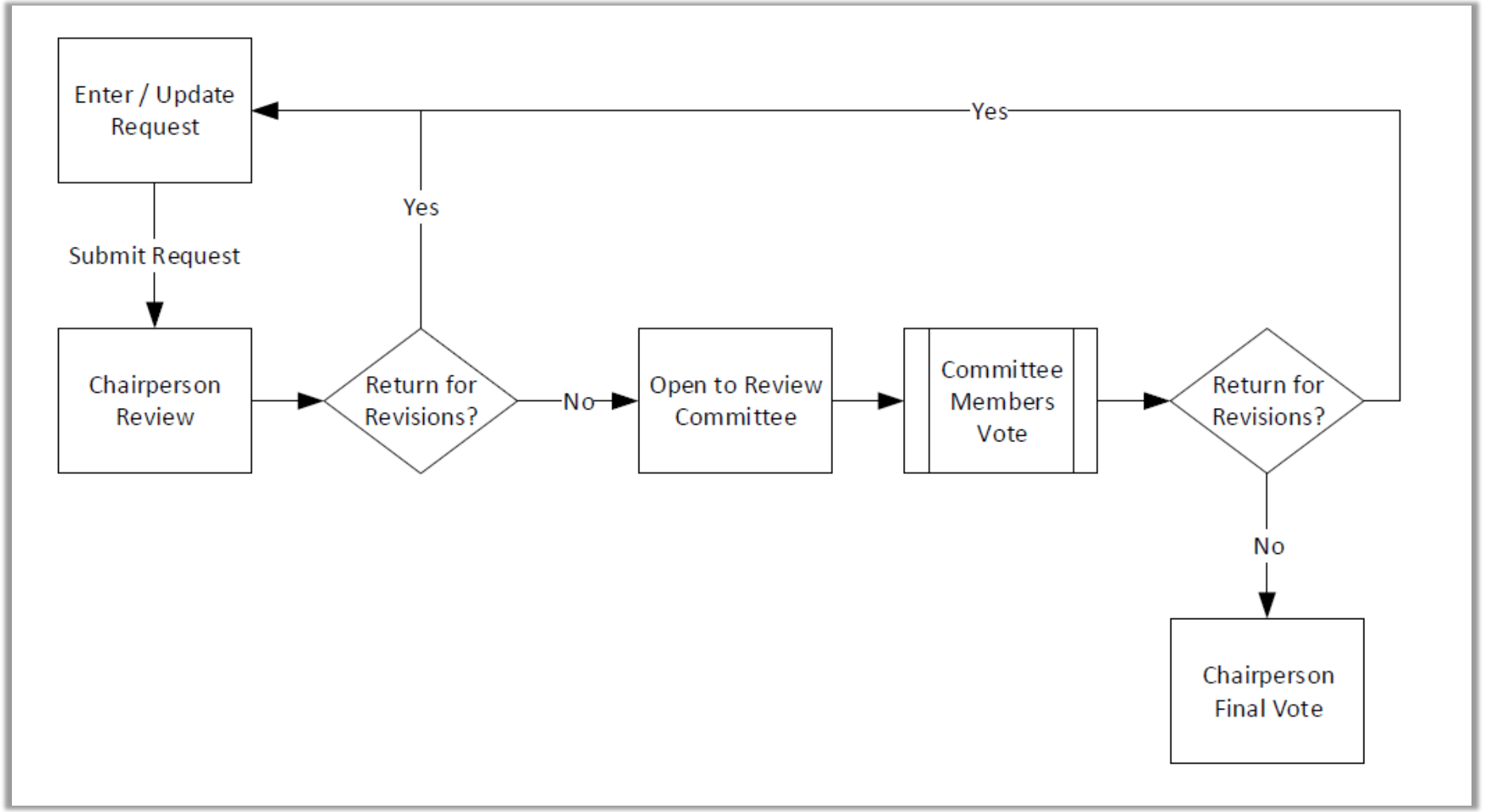
Request Form



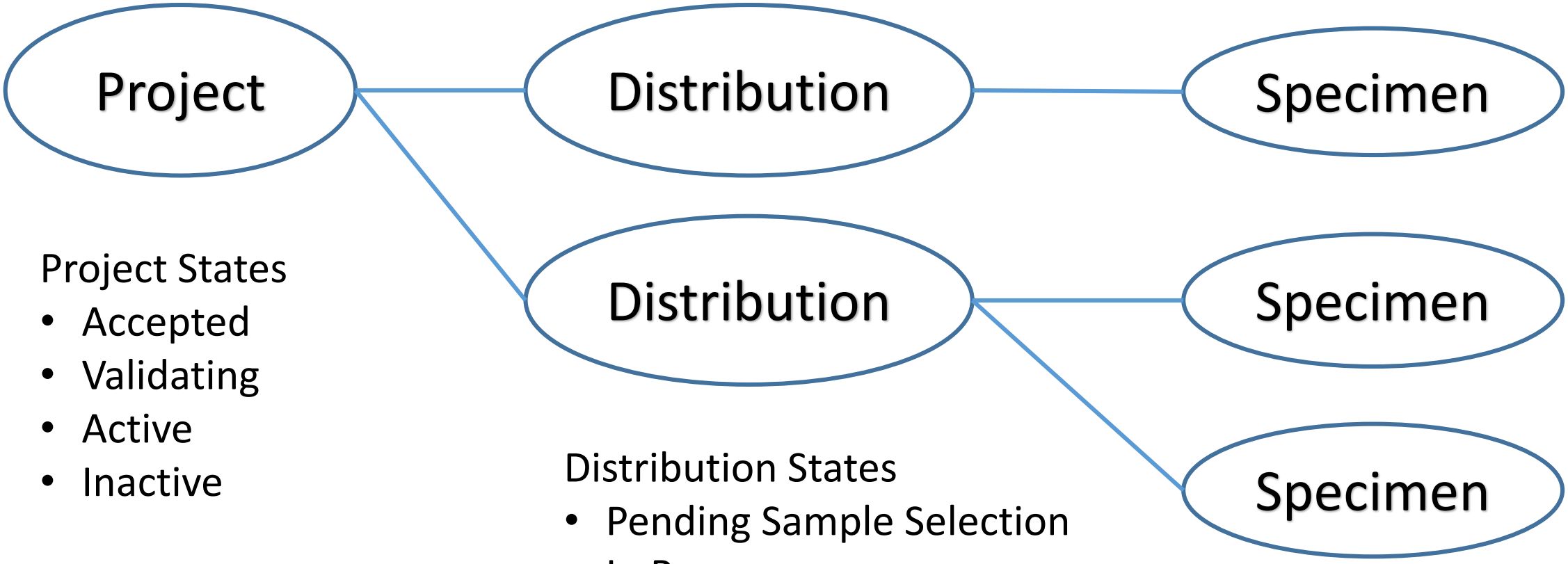




Request Committee Review Process



Accepted External Request = Internal Project

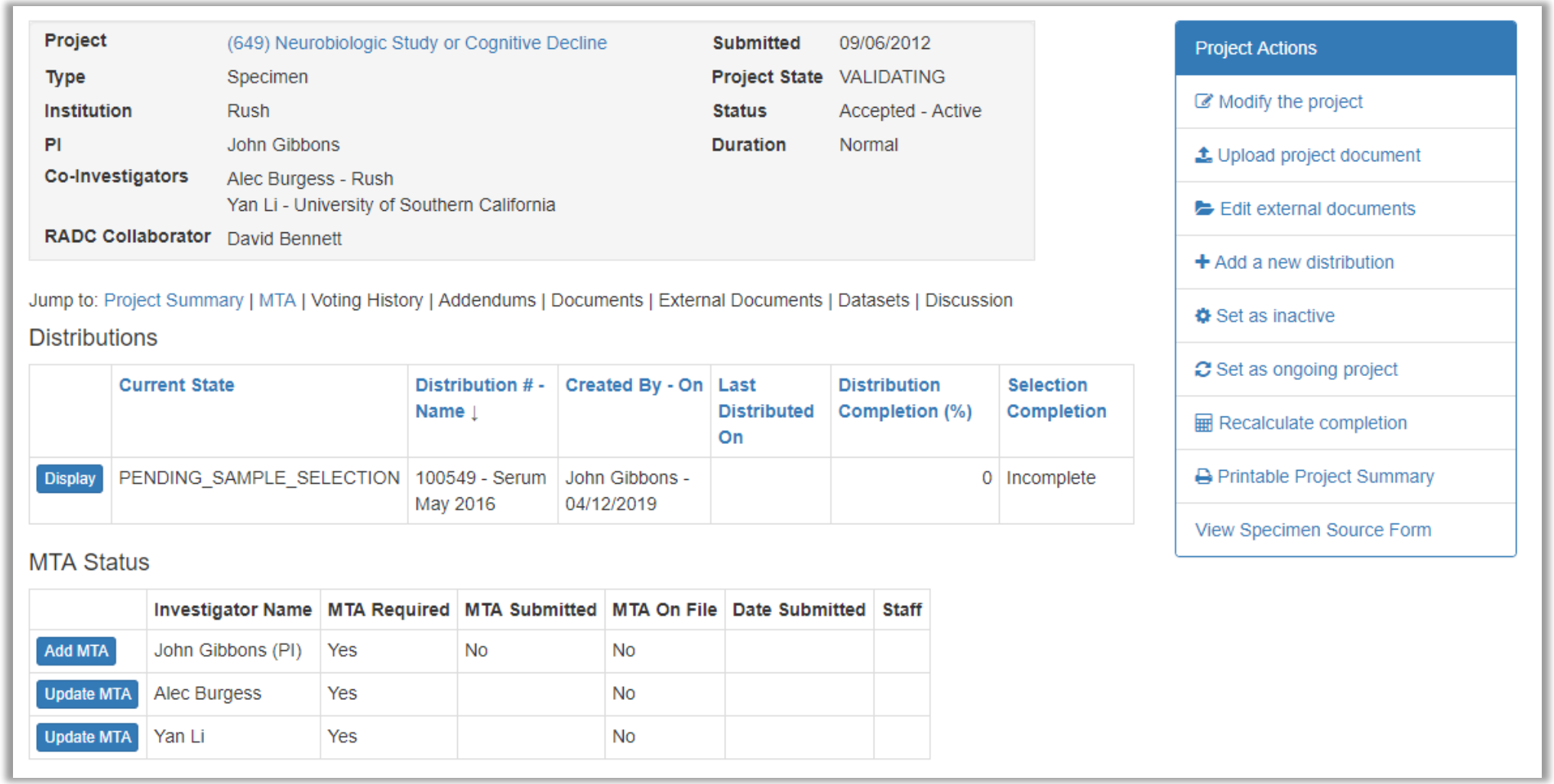


Project States
- Accepted
- Validating
- Active
- Inactive

Distribution States
- Pending Sample Selection
- In Progress
- On Hold
- Cancelled
- Complete

% Complete

Internal Tracking – MTA & DUA, Distributions



Project	(649) Neurobiologic Study or Cognitive Decline	**Submitted**	09/06/2012
Type	Specimen	**Project State**	VALIDATING
Institution	Rush	**Status**	Accepted - Active
PI	John Gibbons	**Duration**	Normal
Co-Investigators	Alec Burgess - Rush Yan Li - University of Southern California		
RADC Collaborator	David Bennett		

Project Actions

- Modify the project
- Upload project document
- Edit external documents
- Add a new distribution
- Set as inactive
- Set as ongoing project
- Recalculate completion
- Printable Project Summary
- View Specimen Source Form

Jump to: Project Summary | MTA | Voting History | Addendums | Documents | External Documents | Datasets | Discussion

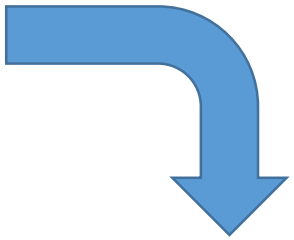
Distributions

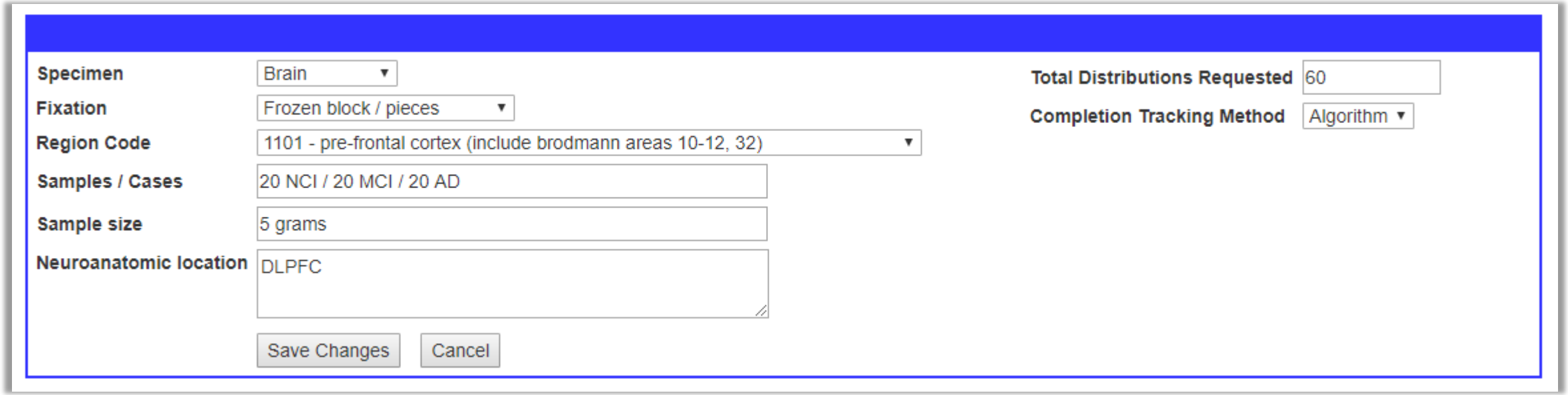
	Current State	Distribution # - Name ↓	Created By - On	Last Distributed On	Distribution Completion (%)	Selection Completion
Display	PENDING_SAMPLE_SELECTION	100549 - Serum May 2016	John Gibbons - 04/12/2019		0	Incomplete

MTA Status

	Investigator Name	MTA Required	MTA Submitted	MTA On File	Date Submitted	Staff
Add MTA	John Gibbons (PI)	Yes	No	No		
Update MTA	Alec Burgess	Yes		No		
Update MTA	Yan Li	Yes		No		

Internal Tracking - Specification of Specimens





Internal Tracking – Specimen/Participant Selection

Many factors to consider:

- Requested inclusion/exclusion criteria
- Protection of “high value” samples
  - Baseline
  - Proximate to death
  - Last collected
- Is there statistical sampling?
- Are there multiple of clinical and pathological diagnoses to consider?

Support team:

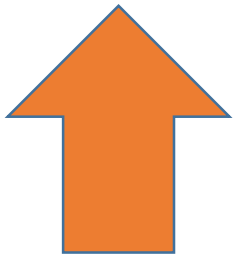
- Data Sharing Coordinator
- Data Management Team
- Statistics Team
- Rush ADC Collaborator

Intake of Requests → Fulfillment → Reporting

- Resource Sharing Hub
- Request Review Process
- DUA & MTA Tracking
- Specimen Selection

- Specimen Distributions
- Plating of Aliquots
- Generation of Data Sets and Codebooks

- Project States
- Tracking Shipments
- Tracking Specimen Distribution



Fulfillment > Aliquot Selection

Serum and Plasma are relatively simple – single aliquots are shared

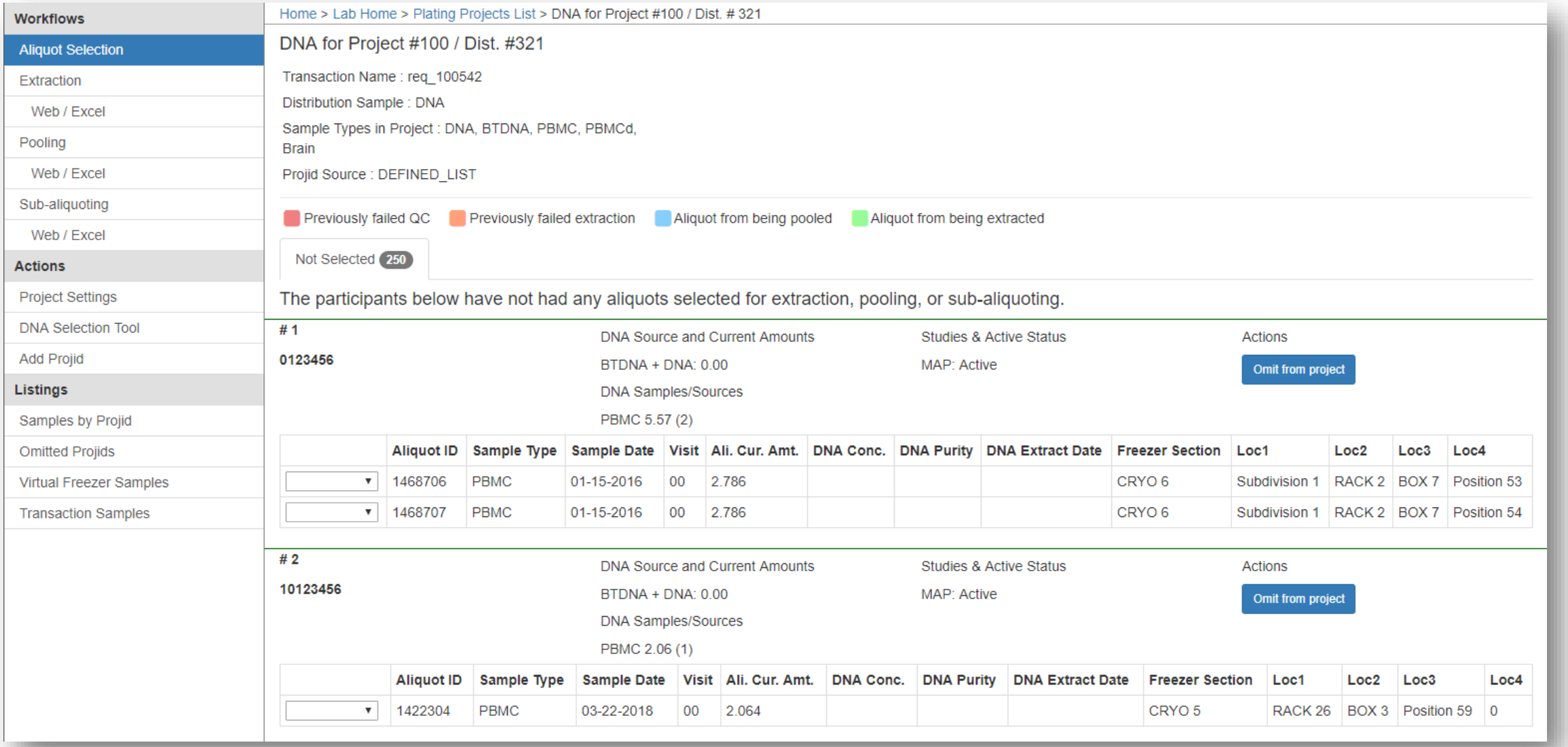
DNA is more complex

- Amount in aliquots varies
- Concentration and purity need to be QC'd
- May need extraction from frozen brain, sediment, or lymphocytes

DNA-specific aliquot workflows

- Pull to QC → Sub-aliquot
- Pooling → QC → Sub-aliquot
- Extraction (frozen brain or PBMC) → QC → Sub-aliquot

Fulfillment > Aliquot Selection

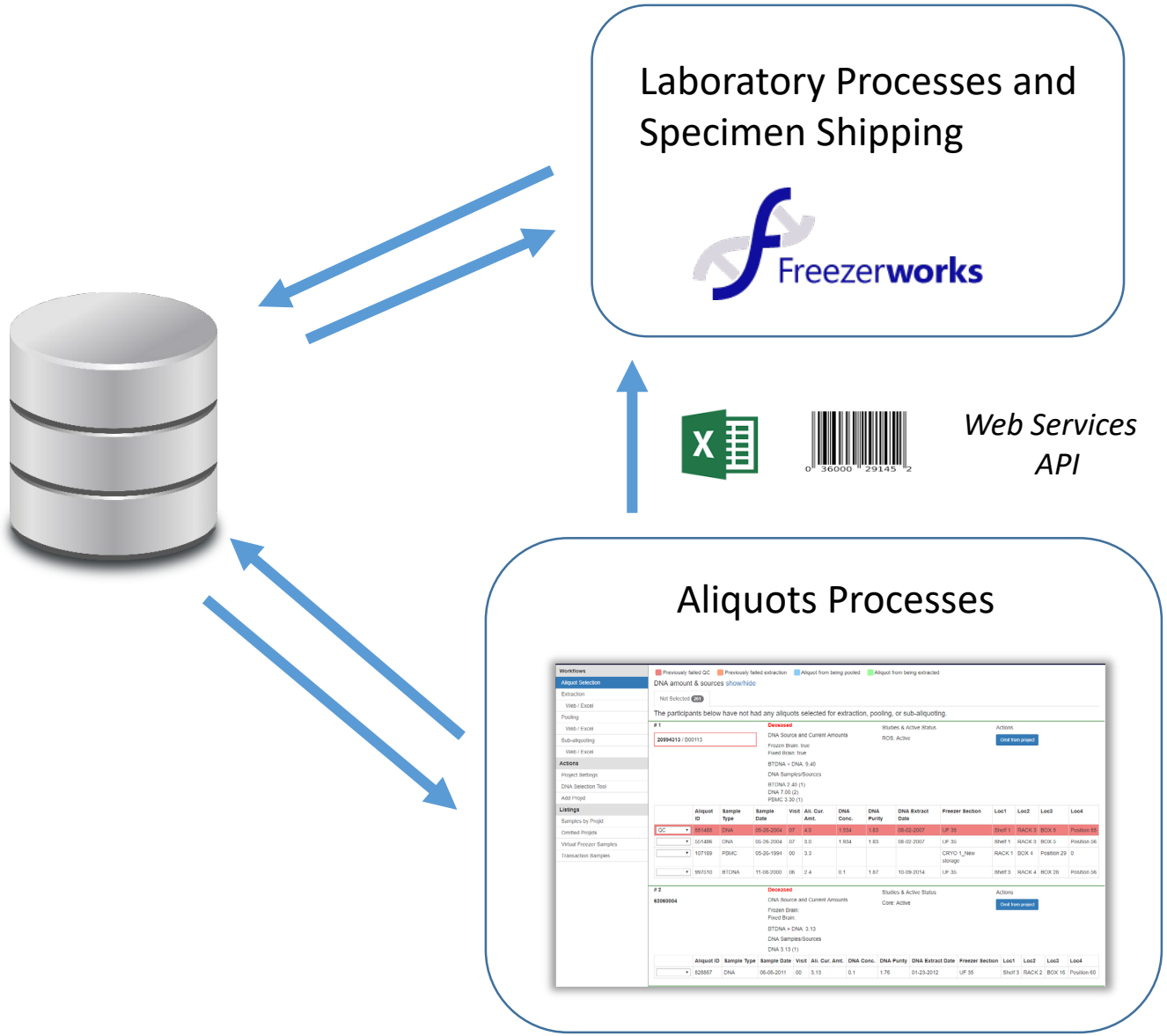


Fulfillment > Tracking Distributions

1. Freezerworks

Tracks lab activities
- Tags aliquots for shipment
- Sub-aliquot process
- Create shipping manifest

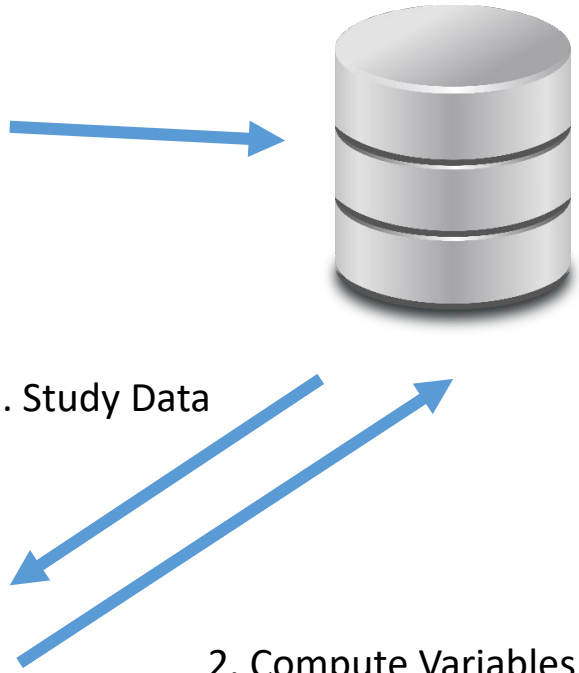
2. Web-based form…
- slides
- Paraffin embedded blocks



Fulfillment Data > Computed Variables System

Study Data Collection

- Laptops – computer assisted interviewing forms
- Web-based forms



Nightly Processing

sas

Central system to document computed variables that are updated nightly.

- 72 database tables
  - 35 longitudinal tables
  - 13 scaled tables
  - 15 table containing special forms of longitudinal variables
    - Baseline
    - Ever
    - Last valid

Fulfillment Data > Computed Variable

Episodic Memory

Property	Value
Name	cogn_ep
External Label	Episodic Memory
Category	Cognition / Domains
Availability	DUA / Hub / External
Type	Longitudinal / Scaled
Short Description	Episodic memory domain – Average or 7 tests
Long Description	See the hub!

Scaled to individual and combinations of studies.

“Cross sectional” in forms:

- Baseline
- Ever
- Last valid
- Cumulative

Fulfillment Data > **Data Sets** of Computed Variables

A custom data set is created for each project linked to each project.

1. Collection of Participants by:
    - Study, Studies, or Substudy
    - Data Criteria
        - Device data
        - Neuropath approved
        - In-vivo MRI
        - Ex-vivo MRI
    - Specific list of identifiers
2. Collection of Variables

Linked to a specific project

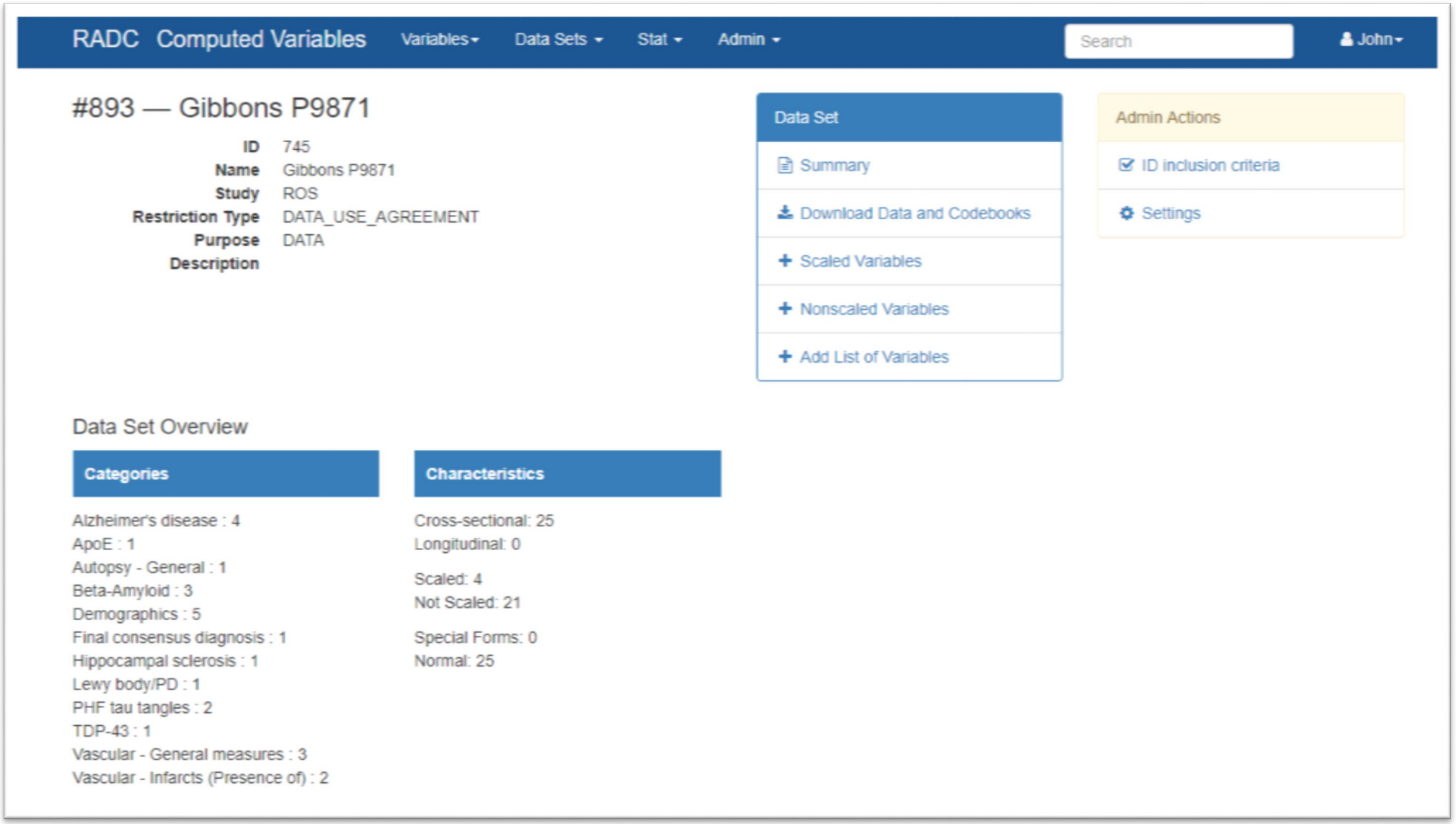
Automatic generation of a custom codebook in a web-based single page format.

Data sets automatically generated as Excel and CSV

Distributions tracked by:

1. Date
2. Participant identifiers
3. List of included variables

Fulfillment Data > Data Sets of Computed Variables



Intake of Requests → Fulfillment → Reporting

- Resource Sharing Hub
- Request Review Process
- DUA & MTA Tracking
- Specimen Selection

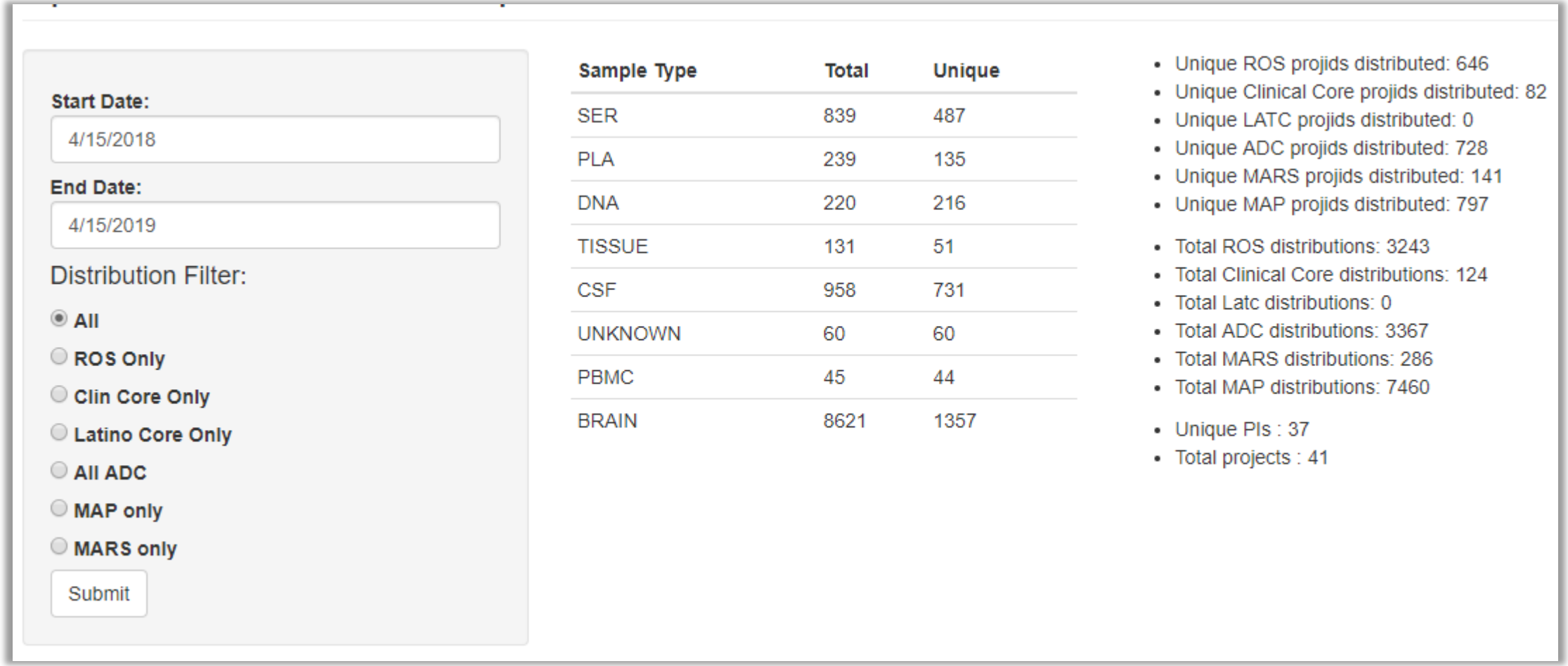
- Specimen Distributions
- Plating of Aliquots
- Generation of Data Sets and Codebooks

- Project States
- Tracking Shipments
- Tracking Specimen Distribution

Project States Report

Project ID	Project State	Primary Investigator	Distribution ID	Distribution State	Sample types
649	VALIDATING	John Gibbons	100549	PENDING	
122	ACTIVE	Yan Li	100188	ON_HOLD	BRAIN [FROZEN BLOCK PIECE] — All MAP frozen brain x 6 regions ◦ VMC: 50 mg ◦ PPUT: 50 mg ◦ MF (BA9/46): 500 mg ◦ IT (BA20): 500 mg ◦ HIP, mid: 500 mg (full cross section) ◦ CALC (BA17): 500 mg
122	ACTIVE	Alec Burgess	100266	ON_HOLD	BRAIN [FROZEN BLOCK PIECE] — All ROSMAP frozen brain ◦ MF (BA9/46) - Region code 1100: 10 mg (2mm punch) x 3
567	VALIDATING	Maylene Liang	100526	PENDING	BRAIN [FROZEN BLOCK PIECE] — 45 (with ffq) - Check in May 2019 for availability ◦ MT (21): 1 gram ◦ MF (9/46): 1 gram ◦ AWS: 1 gram ◦ CBM: 1 gram
1042	ACTIVE	John Gibbons	100300	IN_PROGRESS	BRAIN [FRESH TISSUE] — 150 prospective MAP/ROS with PMI ◦ Prefrontal cortex (BA10) - Region code 1101 replaced MF (BA9/46) - Region code 1100 in October 2015. Region changed back to Midfrontal gyrus in April 2016.: 1.5 x 1.5 x 1.0 cm ◦ AWS. Replaced anterior corpus callosum, region code 2078 8/2016 rej.:

Specimen Distribution Report



Start Date:

4/15/2018

End Date:

4/15/2019

Distribution Filter:

- (•) **All**
- ( ) **ROS Only**
- ( ) **Clin Core Only**
- ( ) **Latino Core Only**
- ( ) **All ADC**
- ( ) **MAP only**
- ( ) **MARS only**

Submit

Sample Type	Total	Unique
SER	839	487
PLA	239	135
DNA	220	216
TISSUE	131	51
CSF	958	731
UNKNOWN	60	60
PBMC	45	44
BRAIN	8621	1357

- Unique ROS projids distributed: 646
- Unique Clinical Core projids distributed: 82
- Unique LATC projids distributed: 0
- Unique ADC projids distributed: 728
- Unique MARS projids distributed: 141
- Unique MAP projids distributed: 797

- Total ROS distributions: 3243
- Total Clinical Core distributions: 124
- Total Latc distributions: 0
- Total ADC distributions: 3367
- Total MARS distributions: 286
- Total MAP distributions: 7460

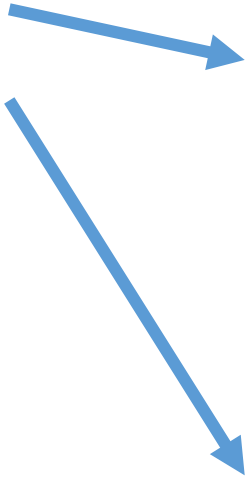
- Unique PIs : 37
- Total projects : 41

Data Distribution Tracking > Ad Hoc Reports

Data Set ID	Distribution ID	Distribution Date
839	1059	5/2/2019

Distributions tracked by:

1. Date
2. Participant identifiers
3. List of included variables



Distribution ID	ProjID
1059	00000136
1059	12399988
1059	34588343
1059	34588734

Distribution ID	Variable Name
1059	cogn_ep
1059	cogn_se
1059	cogn_global
1059	braak
1059	cerad

Team

David Bennett MD, Director
Annie Barz, ADC Administrator

John Gibbons, Data Manager
Alec Burgess, Data Programmer
Yan Li, Data Programmer
Tiffany Wu, Data Programmer

Greg Klein, Data Sharing Coordinator
Maylene Liang, Data Documentation Specialist

Ryan Johnson, Laboratory Manager, Pathologists' Assistant
Karen Skish, Laboratory Manager, Pathologists' Assistant
Srabani Mondal, Manager RADC Lab