BRICS | Biomedical Research Informatics Computing System

# BRICS Demo
## Query Tool and APIs

### Biomedical Research Informatics Computing System (BRICS)

**April 11th, 2024**


CiSTAR

cdRNS
Common Data Repository for Nursing Science

National Ophthalmic Disease Genotyping Network
eyeGENE®

MTBI²
Uniformed Services University

PDBP
Parkinson's Disease Biomarkers Program

National Institute of Neurological Disorders and Stroke

FITBIR
Federal Interagency Traumatic Brain Injury Research
INFORMATICS SYSTEM

# Logistics

| Audio/Video | Please keep your microphone muted |
|---|---|
| **Recording** | • Today's session will be recorded<br>• Will be posted on the BRICS website: https://brics.cit.nih.gov/demo |
| **Questions & Comments** | • We encourage your participation today<br>• Please use the chat for questions & comments. The chat will be monitored throughout today's demo.<br>• There will also be time after each speaker and at the end of the demo to ask live questions. |

# Agenda

| Time | Topic | Speaker(s) |
|---|---|---|
| 9:00 AM-9:10 AM | Introduction and Query Tool Overview | Dr. Matthew McAuliffe |
| 9:10 AM-9:25 AM | Demo on NIA Use Case for Query Tool | Dr. Saba Al-Sayouri<br>Data Scholar |
| 9:25 AM-9:40 AM | Query Tool APIs and Visualization | Dr. Alexandra Bokinsky |
| 9:40 AM-9:55 AM | BRICS APIs and Analysis | Dr. Maria Bagonis<br>Data Scholar |
| 9:55 AM-10:00 AM | Questions & Closing Remarks | Dr. Matthew McAuliffe |

BRICS | Biomedical Research Informatics Computing System

# BRICS - Brief Introduction

Matthew J. McAuliffe, PhD (Biomedical Engineering)
Chief, Scientific Applications Services (SAS), CIT

Home | BRICS (nih.gov)
https://brics.cit.nih.gov

CiSTAR

cdRNS
Common Data Repository for Nursing Science

National Ophthalmic Disease Genotyping Network
eyeGENE®

MTBI²
Uniformed Services University

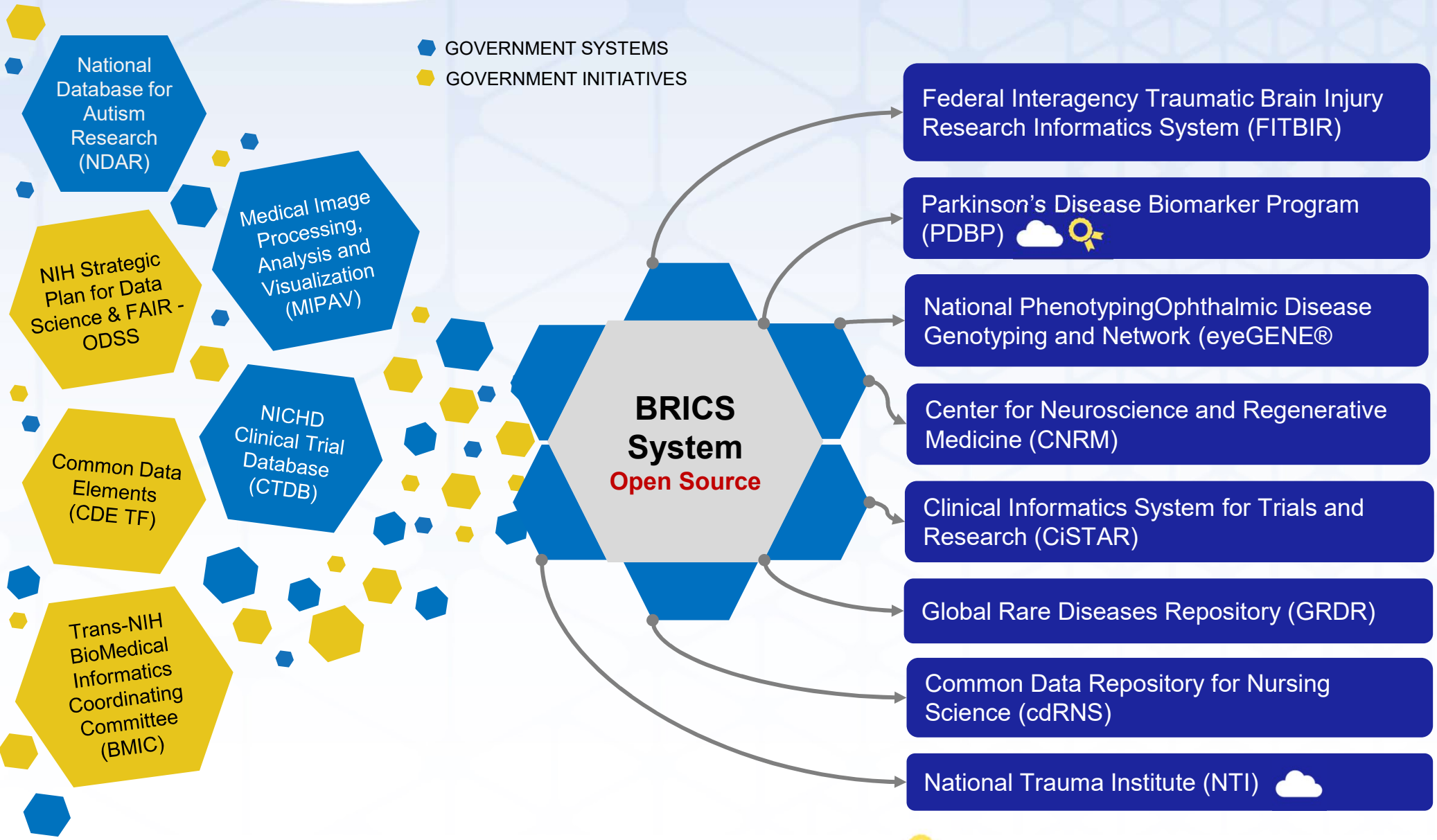
PDBP
Parkinson's Disease Biomarkers Program

National Institute of Neurological Disorders and Stroke

FITBIR
Federal Interagency Traumatic Brain Injury Research
INFORMATICS SYSTEM

BRICS | Biomedical Research Informatics Computing System

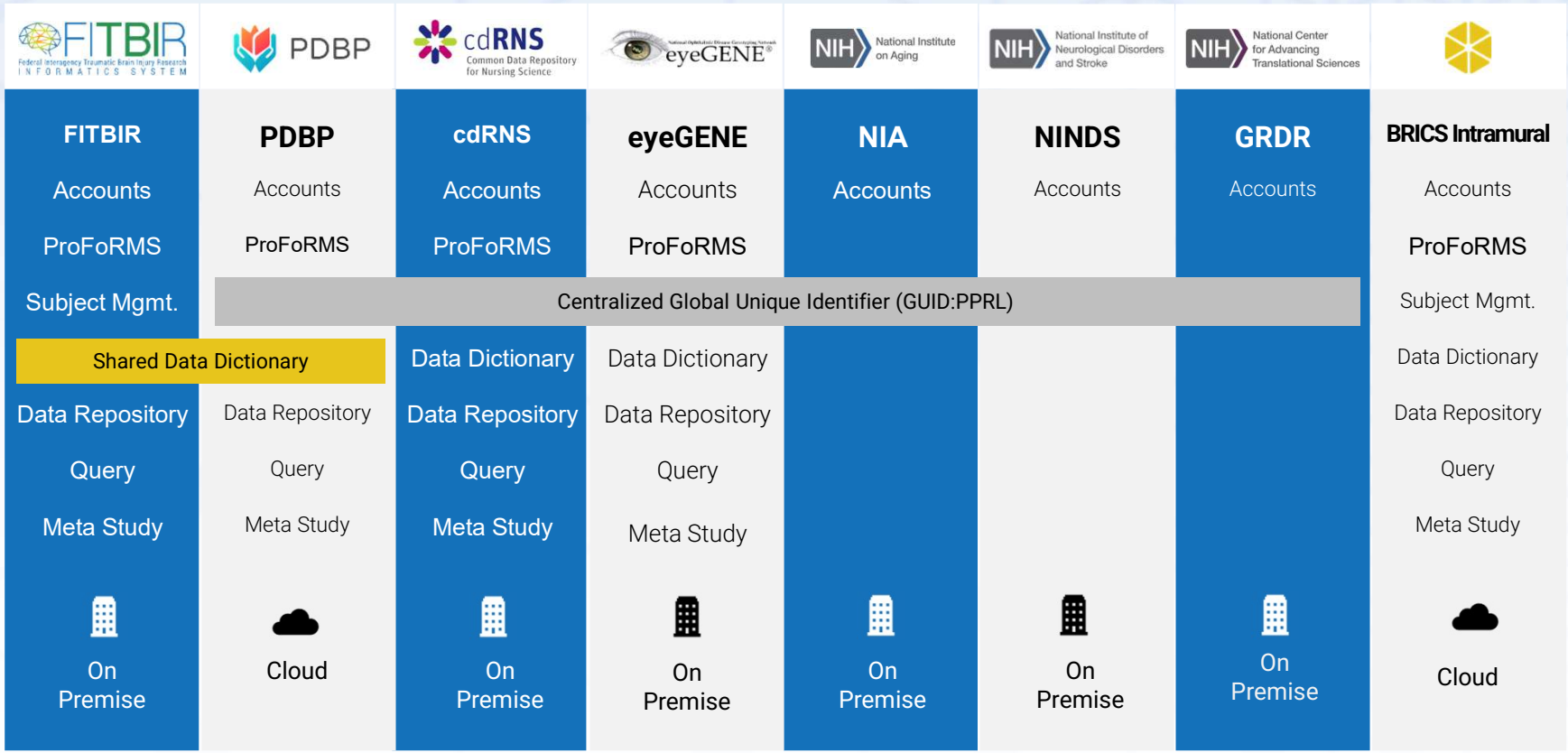
# Building from existing projects




14th Annual Excellence.gov Awards Overall winner

BRICS | Biomedical Research Informatics Computing System

# BRICS Mesh And Fabric




**Datatypes:** Phenotypic, Imaging, and Omics

BRICS | Biomedical Research Informatics Computing System

# Upcoming Releases

**Spiderman Release: (June** 2024)

- Real time saving of locked ProFoRMS data to the repository (Migrating to Mongo DB from PostGres)
- Submission Tool | Globus Integration
- UI Enhancements for the Accounts, Meta Study and Query Tool modules
- Tech library updates as needed
- Various UI and feature enhancements across the BRICS modules

BRICS | Biomedical Research Informatics Computing System

# BRICS Tools | Plug & Play Components for the Full Research Lifecycle

BRICS offers researchers a secure platform and a suite of web-based and downloadable tools that can be shared across disease categories or deployed and branded independently, depending on the needs of your program.

**Data Mapping & Transformation**

Tool that translates data into CDEs used by BRICS to prepare for validation

**Data Repository**

Functionality to define and manage studies and contribute or store data

**Meta Study**

Workspace that aggregates data and metadata across studies for reference

**Data Dictionary**

Intelligent clinical research data dictionary that supports cross-system exchange, CDEs

**ProFoRMS**

Module for electronic data capture (EDC), subject management and scheduling, etc, (21 CFR part 11 compliant)

**Global Unique Identifier**

Cross-study Privacy Preserving Record Linkage (PPRL) system

*Continues…*

BRICS | Biomedical Research Informatics Computing System

### Query

Tool enabling filtering of submitted data using data elements and form structures

### Clinical Trials Management System (CTMS)

Management of Clinical Trials by enabling insight into trial performance.

### BRICS Imaging Tools

Enables quantitative analysis and viewing of medical images, such as PET, MRI, CT, or microscopy.

### Forum

Discussion board for account users for posting messages, interacting with each other, and discussing various topics

### Account Management

Create, approve, and manage user accounts. (RAS enabled)

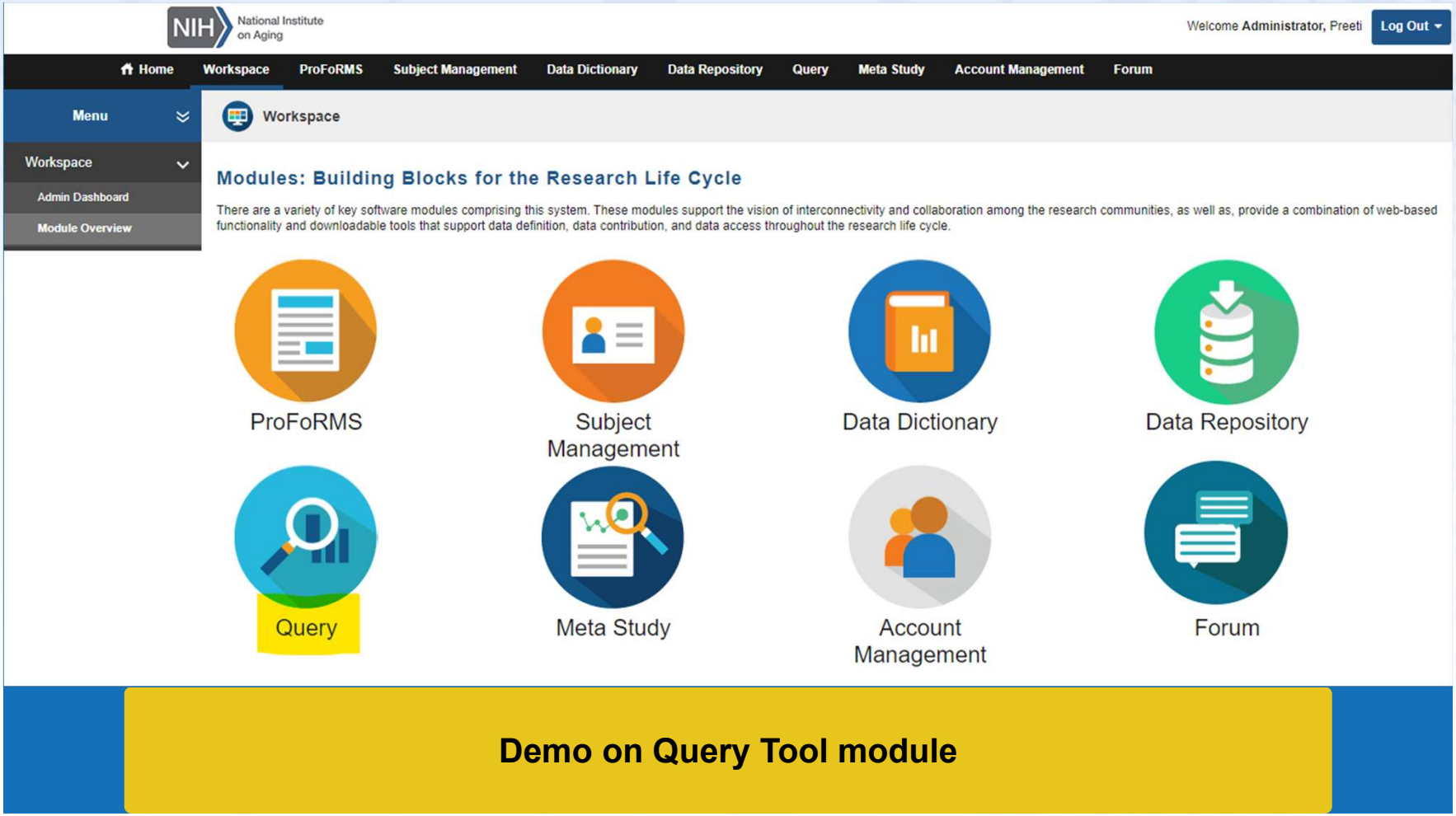
### InET

Application for assigning new training, tracking current training, licenses and compliance.

### STAMS: Specimen Tracking And Management System

Provides the path to container storage locations (e.g. freezer) that guides the user to the specific container and sample.

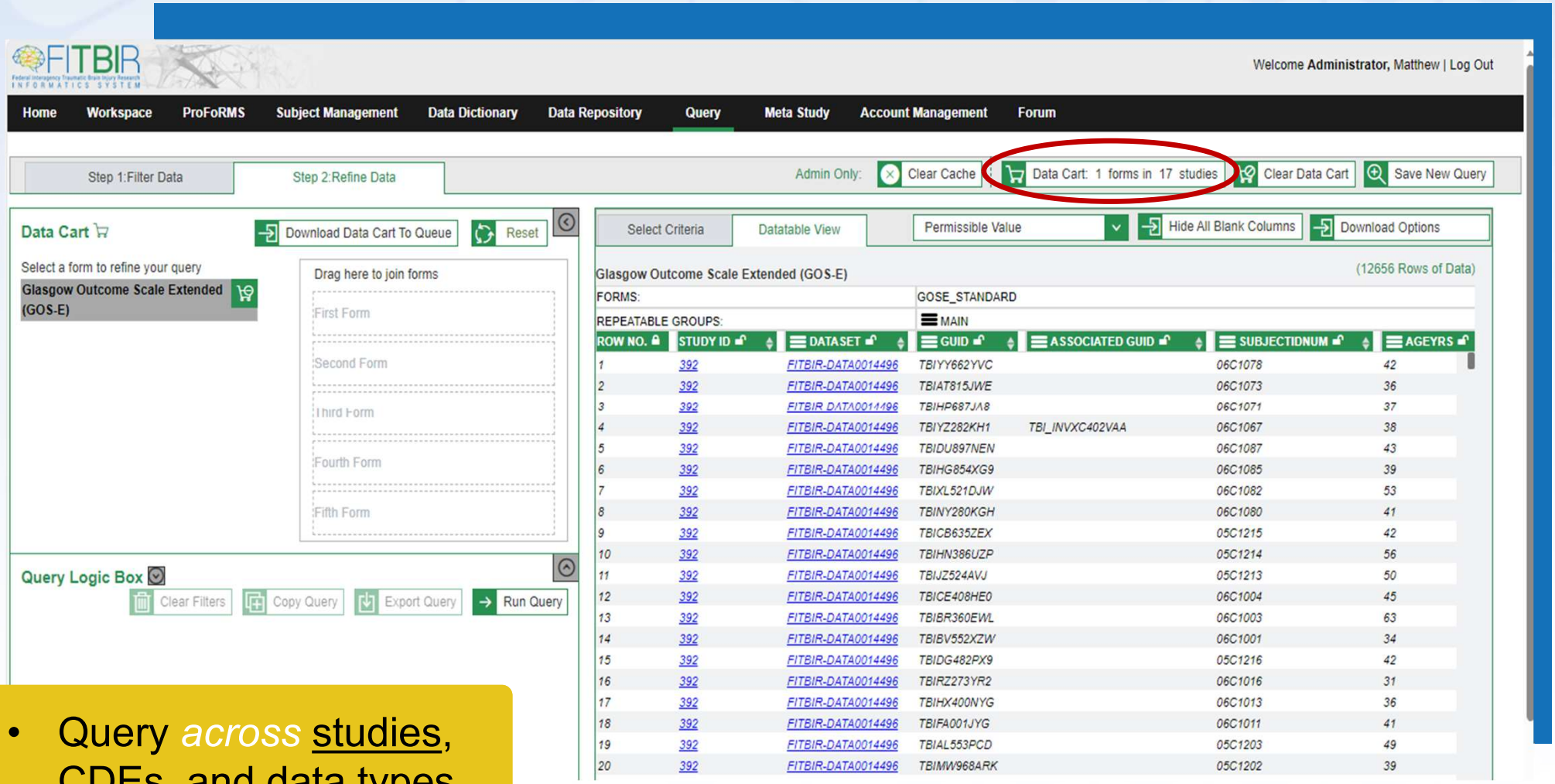
# BRICS Modules




**Demo on Query Tool module**

BRICS | Biomedical Research Informatics Computing System

# Query Tool




- Query *across* studies, CDEs, and data types
- API

BRICS | Biomedical Research Informatics Computing System

# API Query Tool
# Programmatic access to the data

*Authenticate RAS (11/2023)*
- *Query*
- *Retrieve data*
- *Analysis*

**Researcher**
**Data Scientist**

**Jupyter Notebooks**
**Python, R,**
**AI/ML libraries**

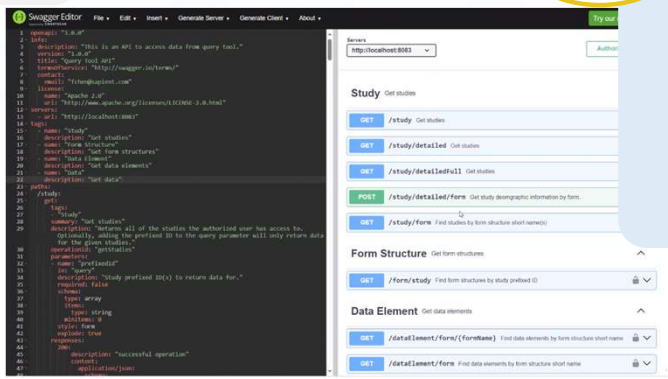
**Rest API**

**Rest API**

**FITBIR Data**

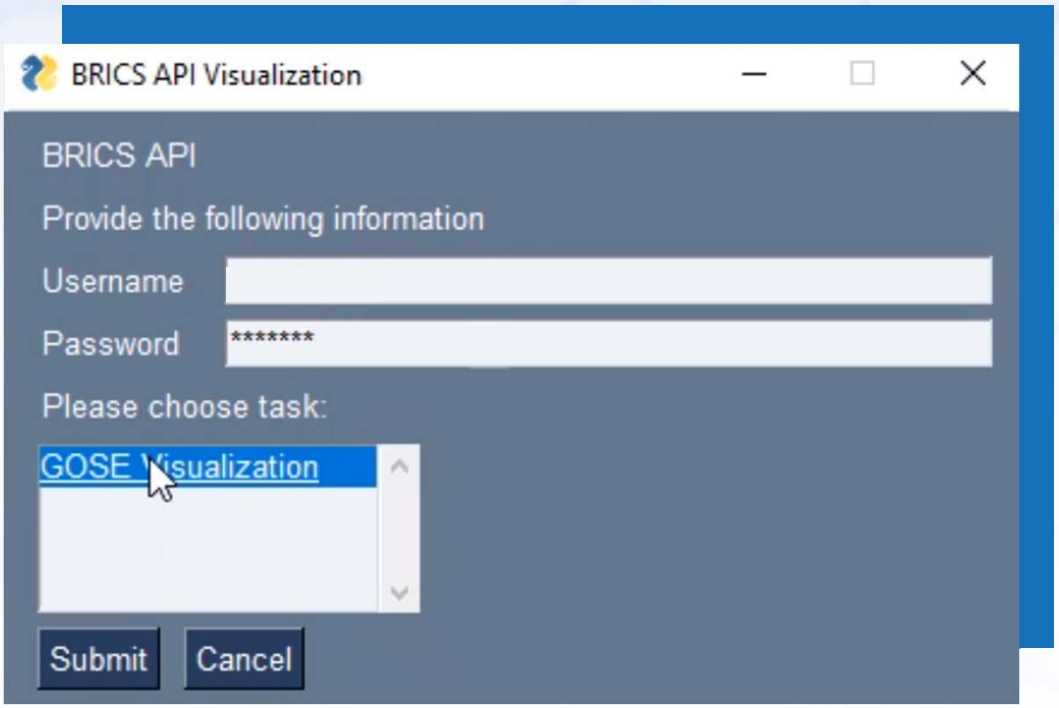
**PDBP Data**

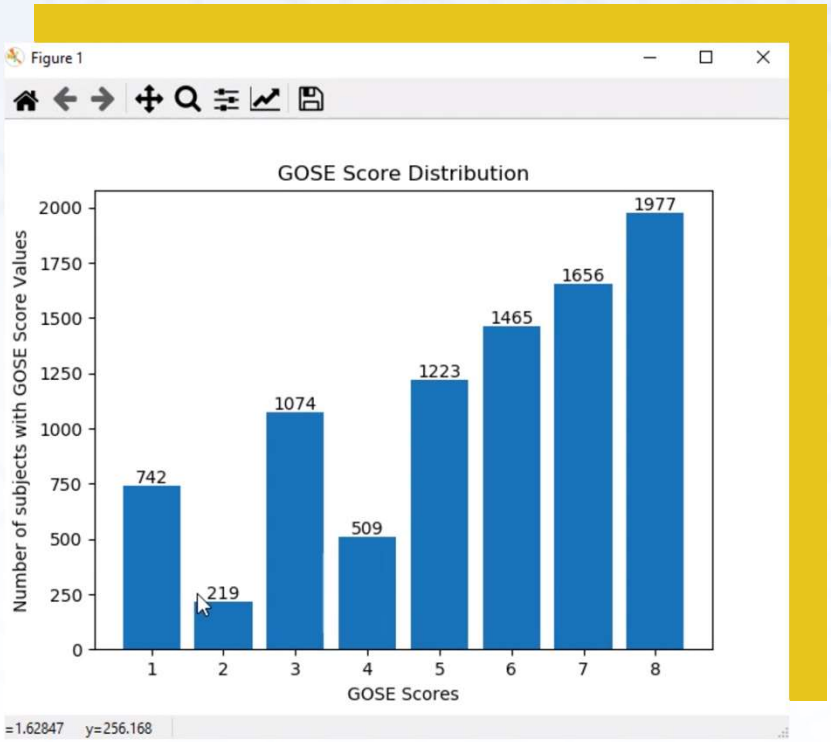
**BRICS Informatics Systems**



# Application GUI – Query/Output

## Python Code



BRICS | Biomedical Research Informatics Computing System

CiSTAR

cdRNS
Common Data Repository for Nursing Science

National Ophthalmic Disease Genotyping Network
eyeGENE®

MTBI²
Uniformed Services University

PDBP
Parkinson's Disease Biomarkers Program

National Institute of Neurological Disorders and Stroke

FITBIR
Federal Interagency Traumatic Brain Injury Research
INFORMATICS SYSTEM

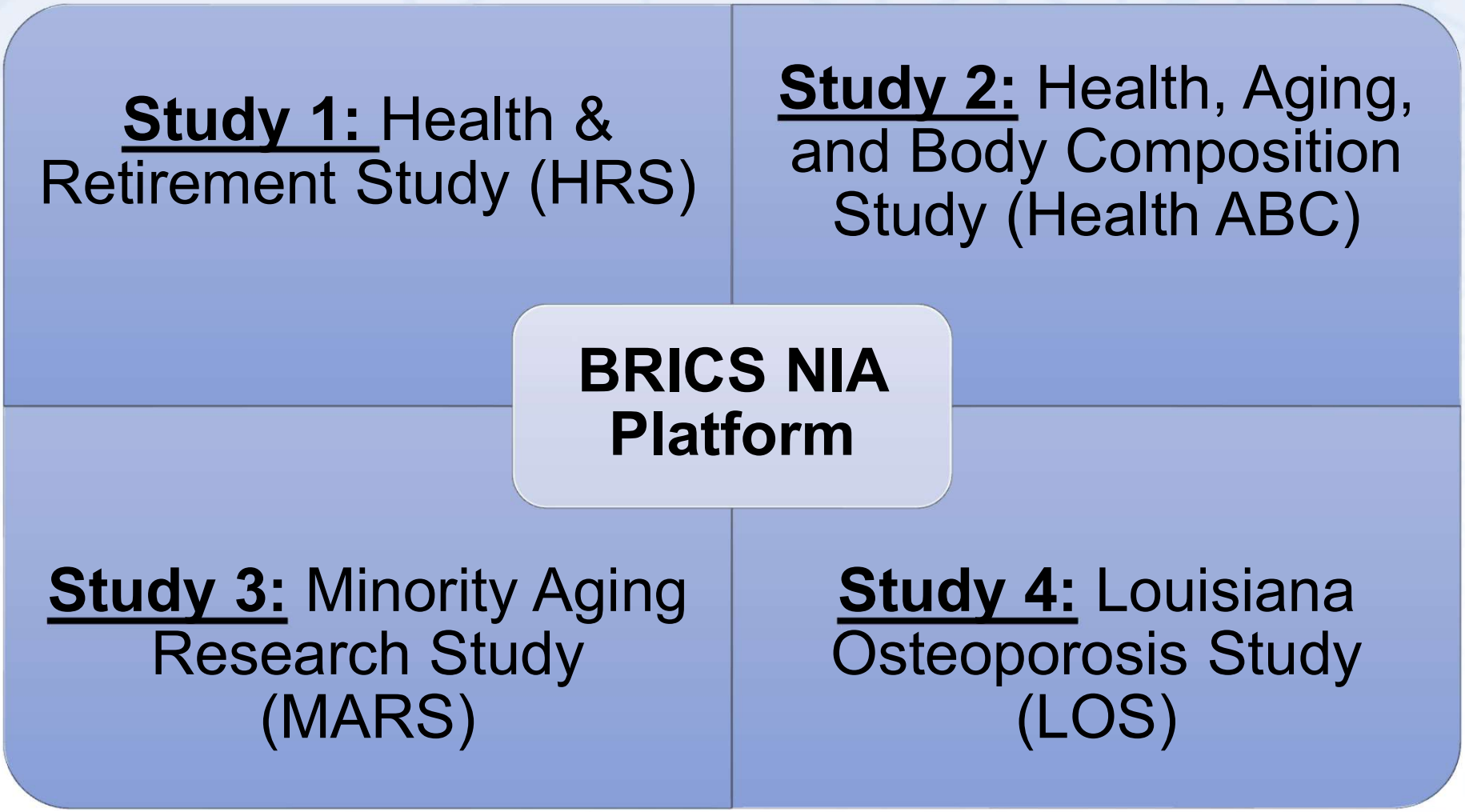
## Biomedical Research Informatics Computing System (BRICS)

# National Institute on Aging (NIA) Pilot Project

## Query Tool Use Case

Dr. Saba Al-Sayouri
Data Scholar

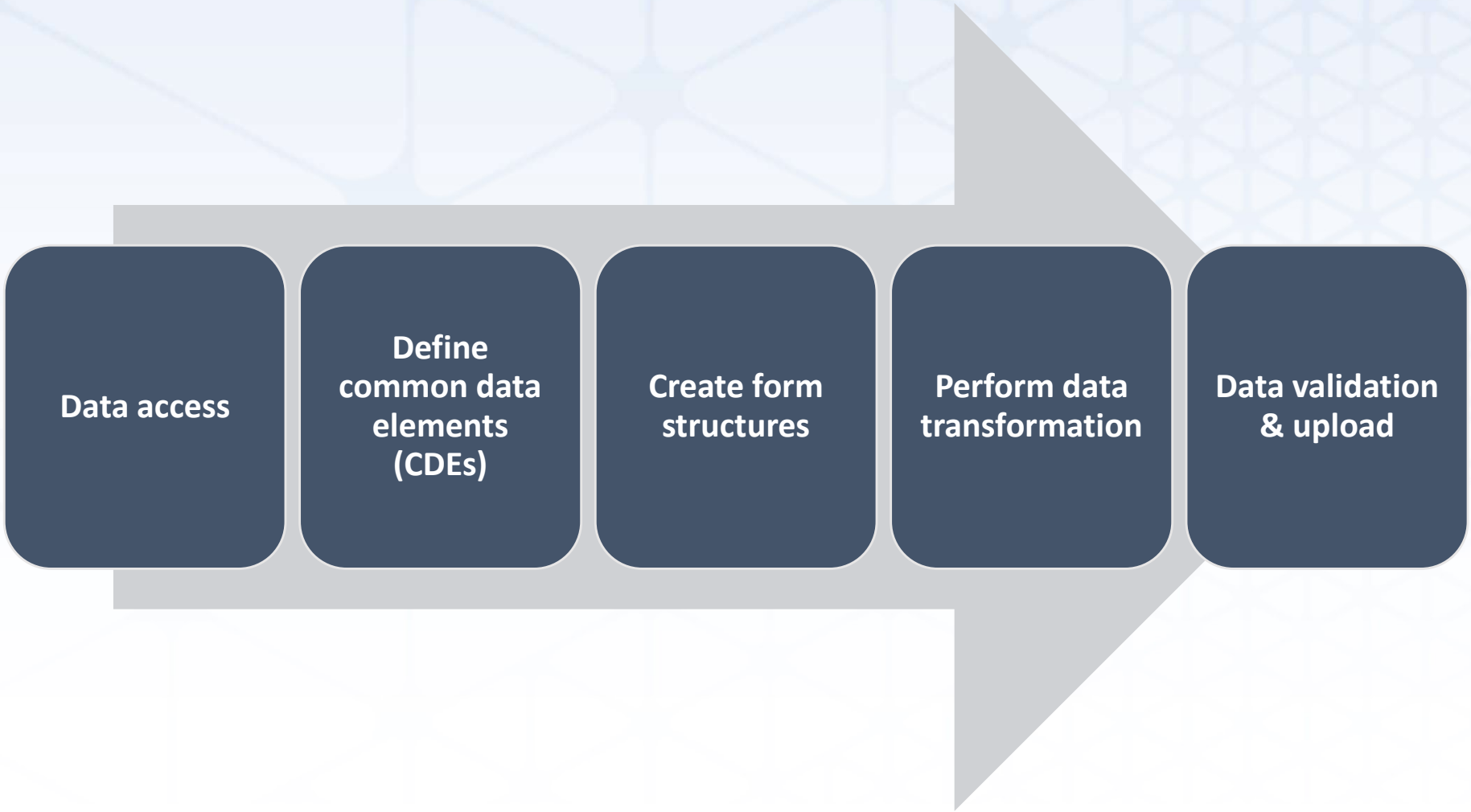
# Studies Overview



# BRICS Added Value - Cohort Discovery

1. Assemble multi-level data collected across lifespan & providing framework for multi-disciplinary work
   - Expedite epidemiological discovery
2. Foster communication among investigators leading aging studies
3. Promote collaborative research projects for topics not easily addressed by single study

# Extract Transform Load (ETL) Process Steps



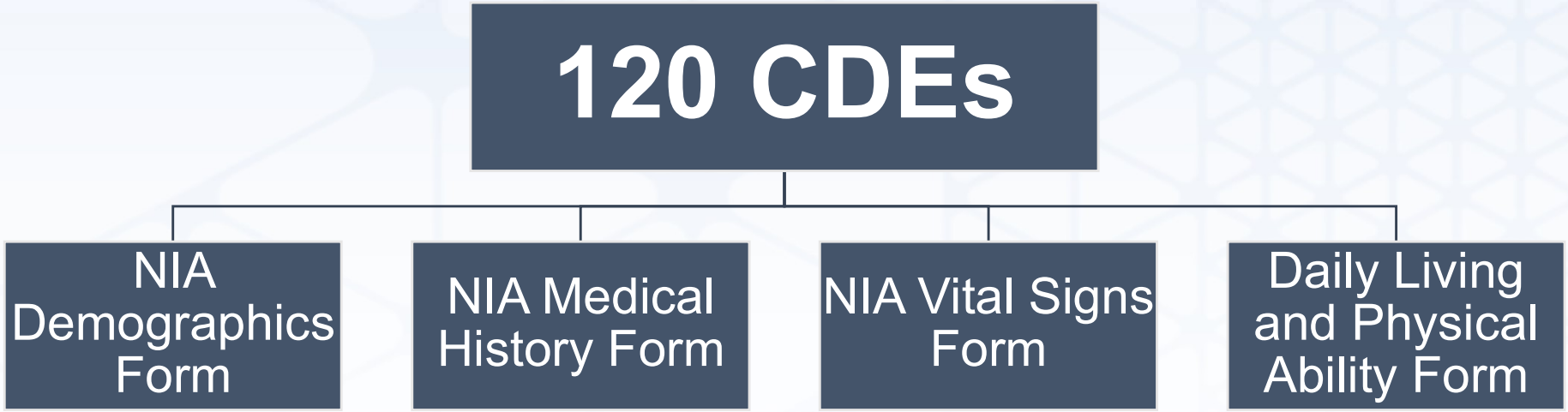
# ETL Process Step 1: CDE Definition

1. Accessed 4 studies data sets
2. Identified ~120 Common Data Elements (CDEs)
   - CDE: variable presents in 2+ studies
   - Iterative process: requires multiple phases of combing through data documentation & dictionaries

| CDE Title | Health ABC | HRS | MARS | LOS |
|---|---|---|---|---|
| Data source | | | | |
| Data source other | | | | |
| General notes | | | | |
| Language of form | | | | |
| Site name | | | | |
| Subject ID | | | | |
| Visit date | | | | |
| Year data collected | | | | |
| Study year | | | | |

Snapshot of Data Inclusion Factsheet

# ETL Process Step 2: Creating Form Structures



# ETL Process Step 3: Data Transformation

- Similar CDEs collected differently across studies
  - Transformed data to fit into BRICS CDE definitions

**Health ABC**

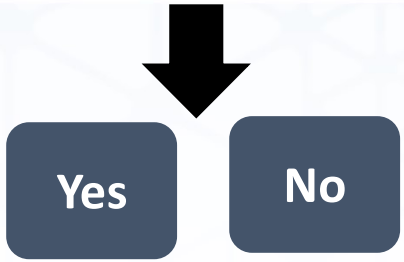
| Variable | General Descriptio | Value labels |
|---|---|---|
| DIFFPP | Difficulty pushing/pulling | 0=No<br>1=Yes |

**HRS**

How difficult is it for you to...

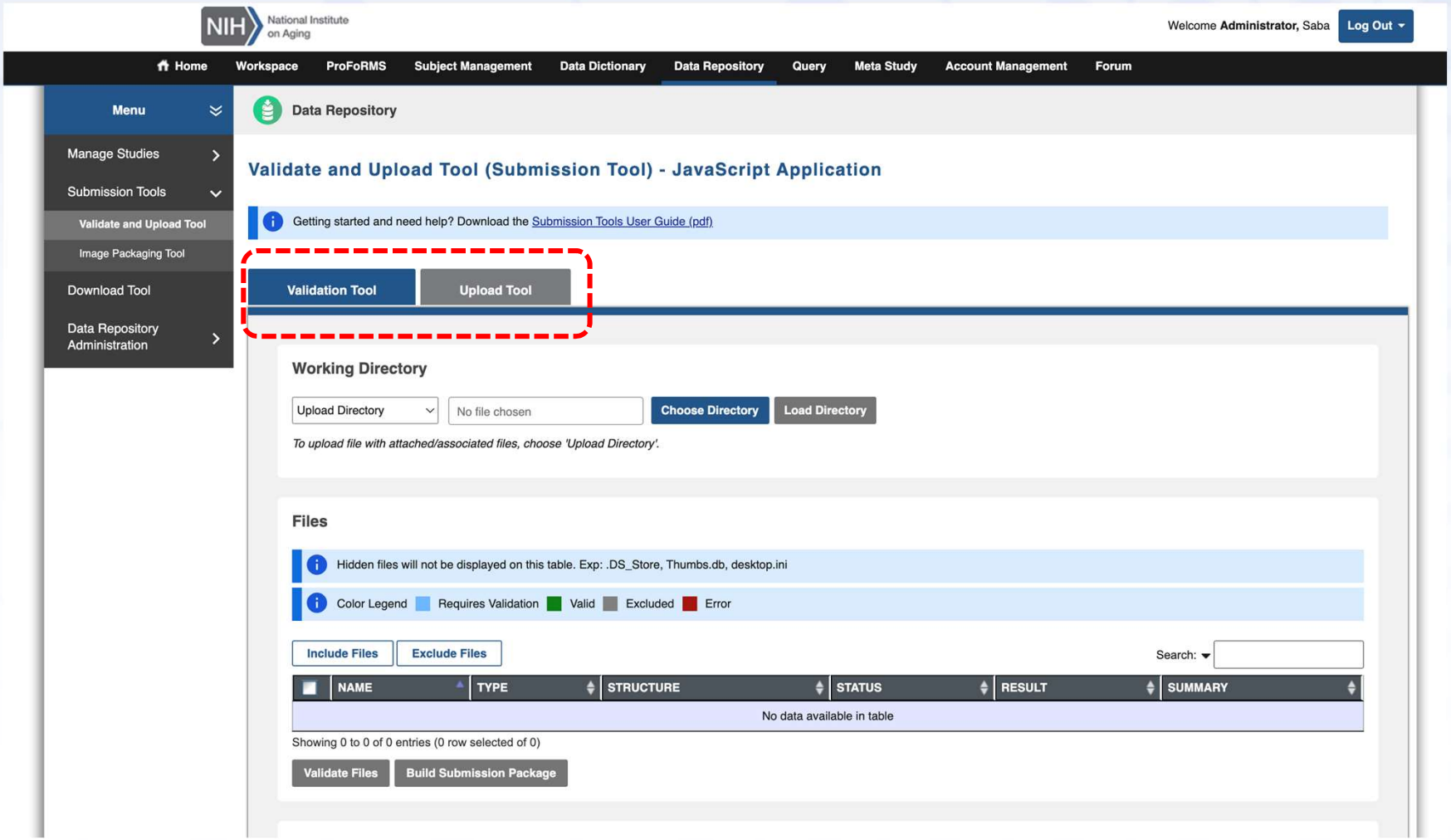
B4r. ...pull or push large objects like a living room chair? [IMPUTED]

| V318 Code | Frequency |
|---|---|
| 1 | 9303 |
| 2 | 1577 |
| 3 | 766 |
| 4 | 843 |
| 6 | 163 |




✅ Harmonized

# ETL Process Step 4: Data Validation & Upload



# Where we Are Now?

- ~120 harmonized CDEs across 4 NIA studies grouped into 4 form structures reflecting 1 year of study data for **16,613** participants
  - Uploaded **all** years available data of MARS study
- We're able to query data across studies, with lots of flexibility in filtering
- We're able to download harmonized data from multiple studies with complete data dictionary

# Future Work

1. Year 1 of HRS full data in process
   - Still in pilot phase, with all data remaining private
   - Ability to run analyses, with all data & documentation in clear, accessible, & easy to use format

2. Working on uploading data for harmonized CDEs for year 2 & across remaining 3 studies (HRS, Health ABC, & LOS)

# Demo

BRICS | Biomedical Research Informatics Computing System

# Questions?

BRICS | Biomedical Research Informatics Computing System

CiSTAR

cdRNS
Common Data Repository for Nursing Science

National Ophthalmic Disease Genotyping Network
eyeGENE®

MTBI²
Uniformed Services University

PDBP
Parkinson's Disease Biomarkers Program

National Institute of Neurological Disorders and Stroke

FITBIR
Federal Interagency Traumatic Brain Injury Research
INFORMATICS SYSTEM

**Biomedical Research Informatics Computing System (BRICS)**

# Query Tool API & Data Visualization

Dr. Alexandra Bokinsky

BRICS | Biomedical Research Informatics Computing System

# API Query Tool
# Programmatic access to the data

*Authenticate RAS (11/2023)*

- *Query*
- *Retrieve data*
- *Analysis*

**Researcher**
**Data Scientist**
**Visualization**

**Jupyter Notebooks**
**Python, R,**
**AI/ML libraries**
**JavaScript**

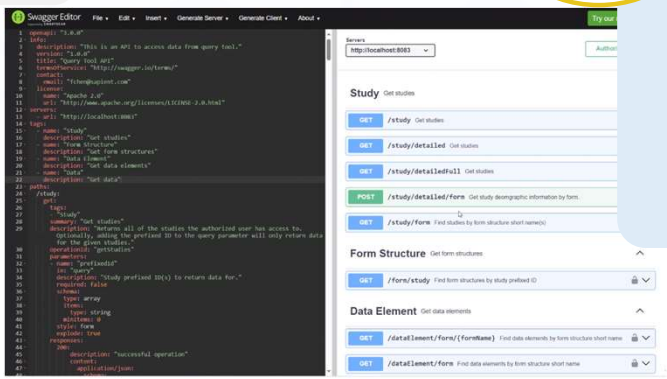
**Rest API**

**Rest API**

**FITBIR Data**

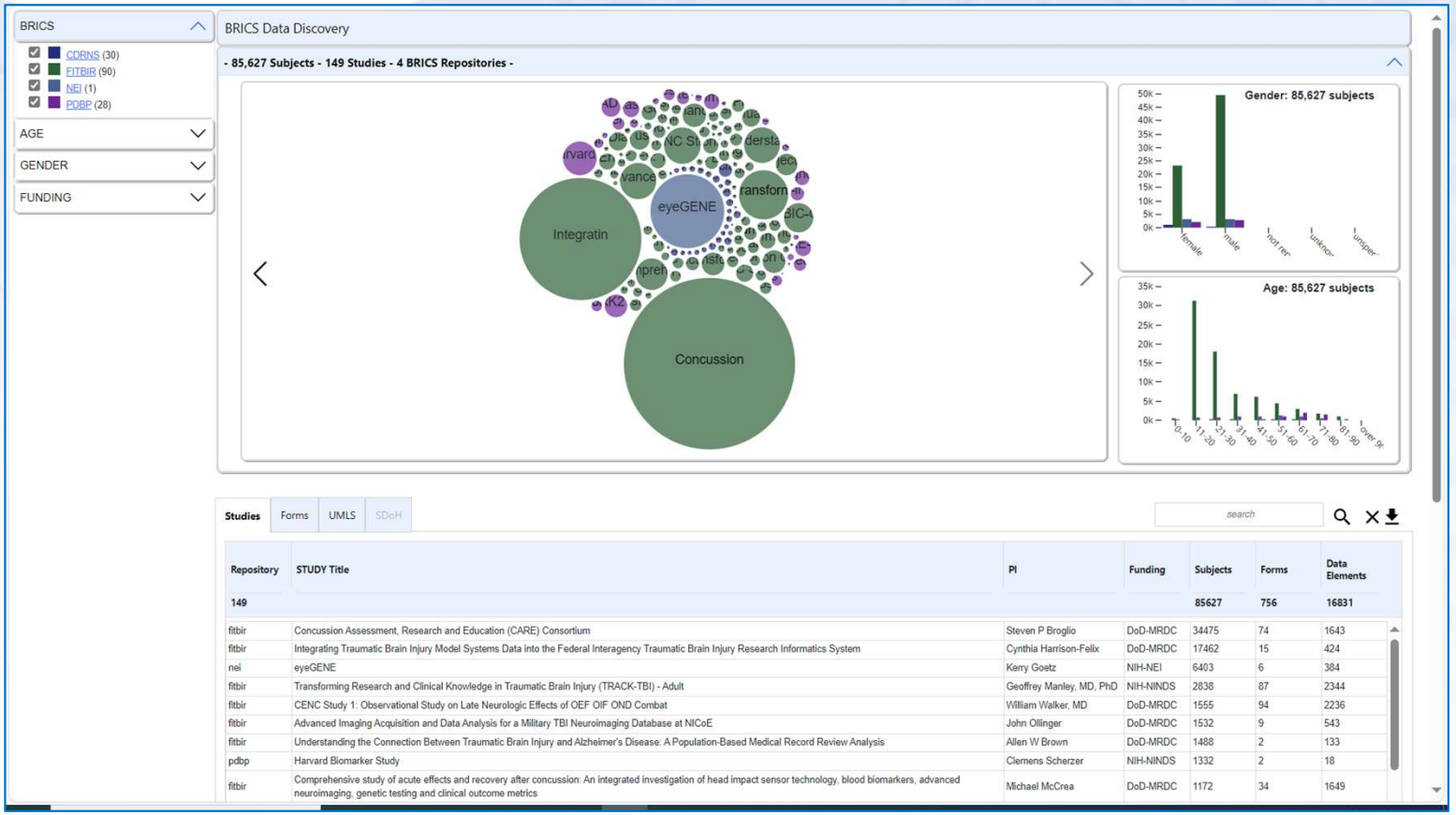
**PDBP Data**

**BRICS Informatics Systems**



BRICS | Biomedical Research Informatics Computing System

# Why would we want to use the API?




- ✓ Programmable access to the Query Tool includes all QT functionality
- ✓ Single program runs on any BRICS instance – same code!
- ✓ Easily build programs to search and integrate data across multiple BRICS instances

BRICS | Biomedical Research Informatics Computing System

# How do we use the API?

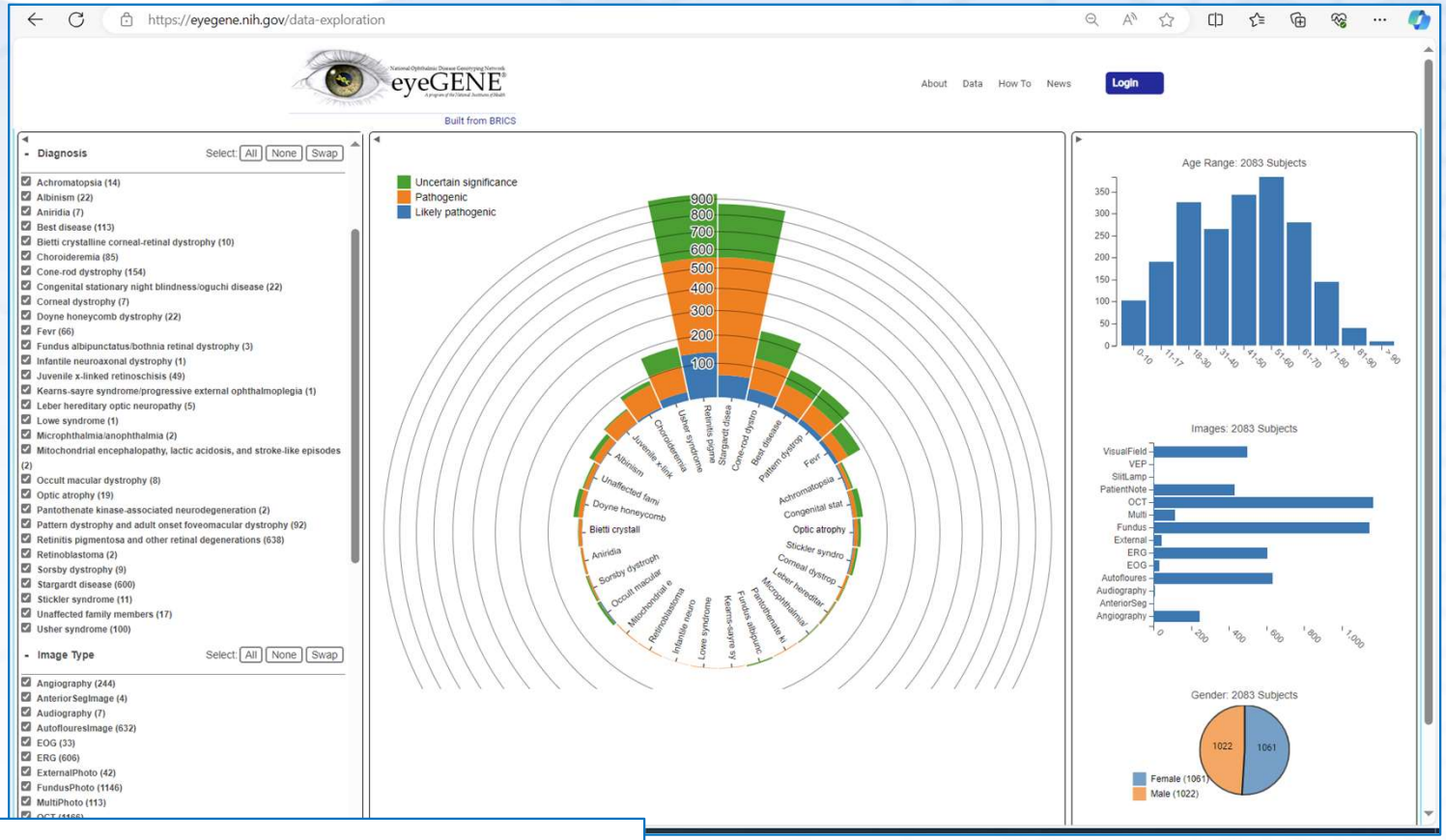

**Each BRICS instance requires a unique RAS authentication token to access the Query API programmatically.**

**API calls are identical across instances – only the token and base url change**

FITBIR: download activity over time by study, data type, and image modality

BRICS | Biomedical Research Informatics Computing System

# How do we use the API?

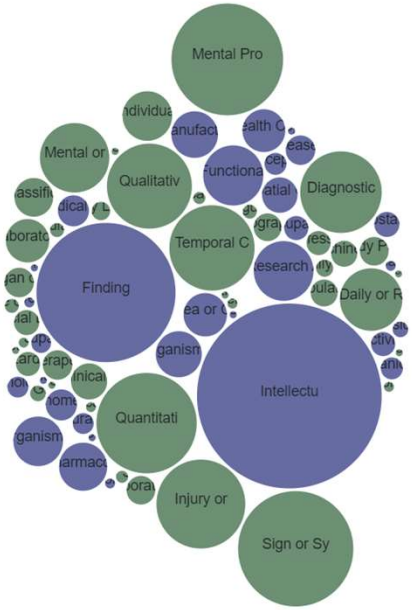



**Log onto the BRICS workspace and navigate to Account Management**

- **Token expires after 30 minutes**
- **Renew token programmatically from the API: /rasauth/user/extendApi**

Diagnostic information, by age, gender, image type, and gene variant

# How do we use the API?




/repository-api/study

**/query-api/form/study?prefixedId=name**

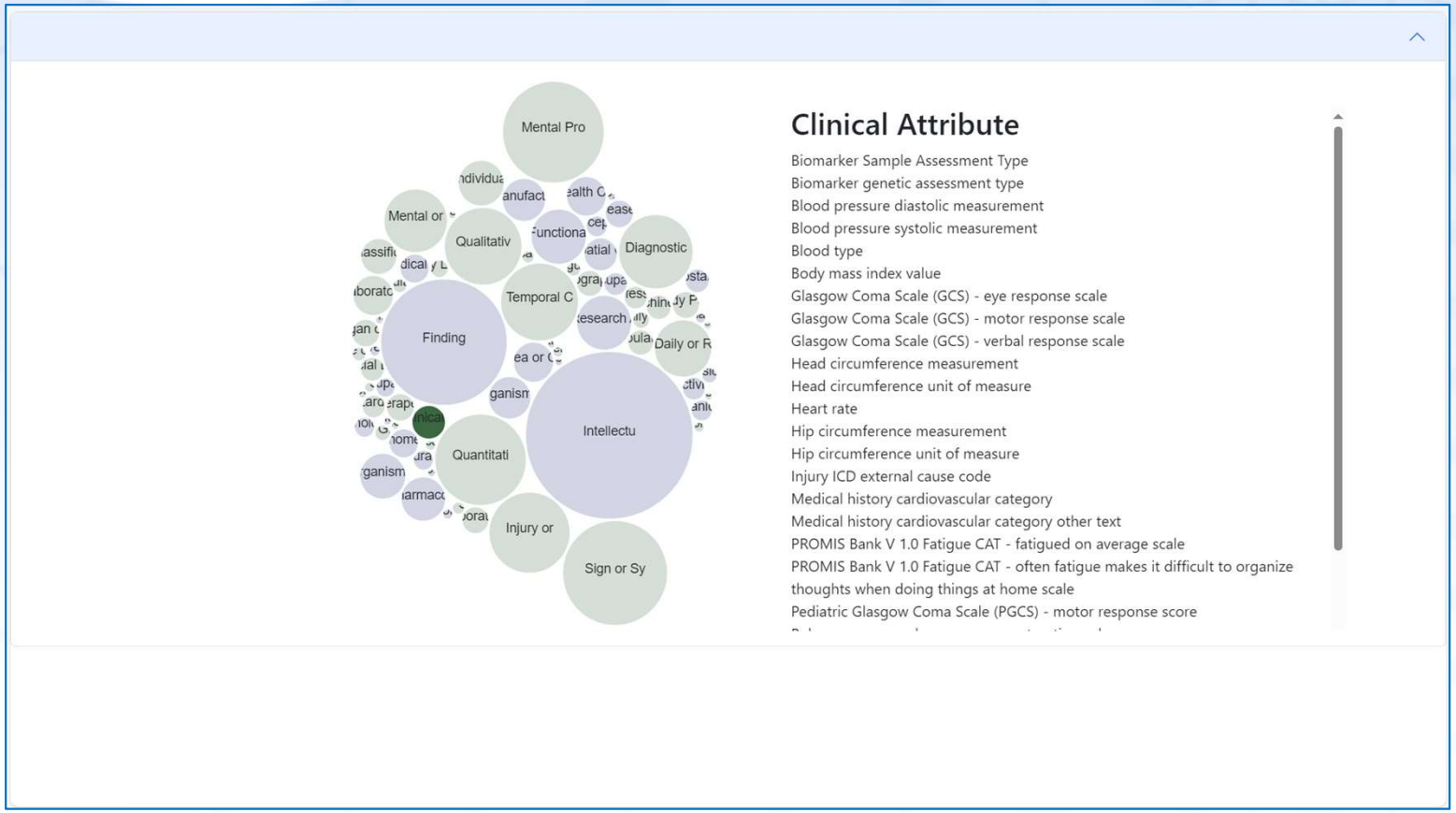
**/query-api/dataElement/form?formName=name**

/dictionary-api/dataElement/getByNames

- Repository-API: List studies available for a given BRICS instance
- **Query-API: List forms for a given study**
- **Query-API: List data elements for a given form**
- Dictionary-API: Return dictionary (including UMLS) information for a given CDE

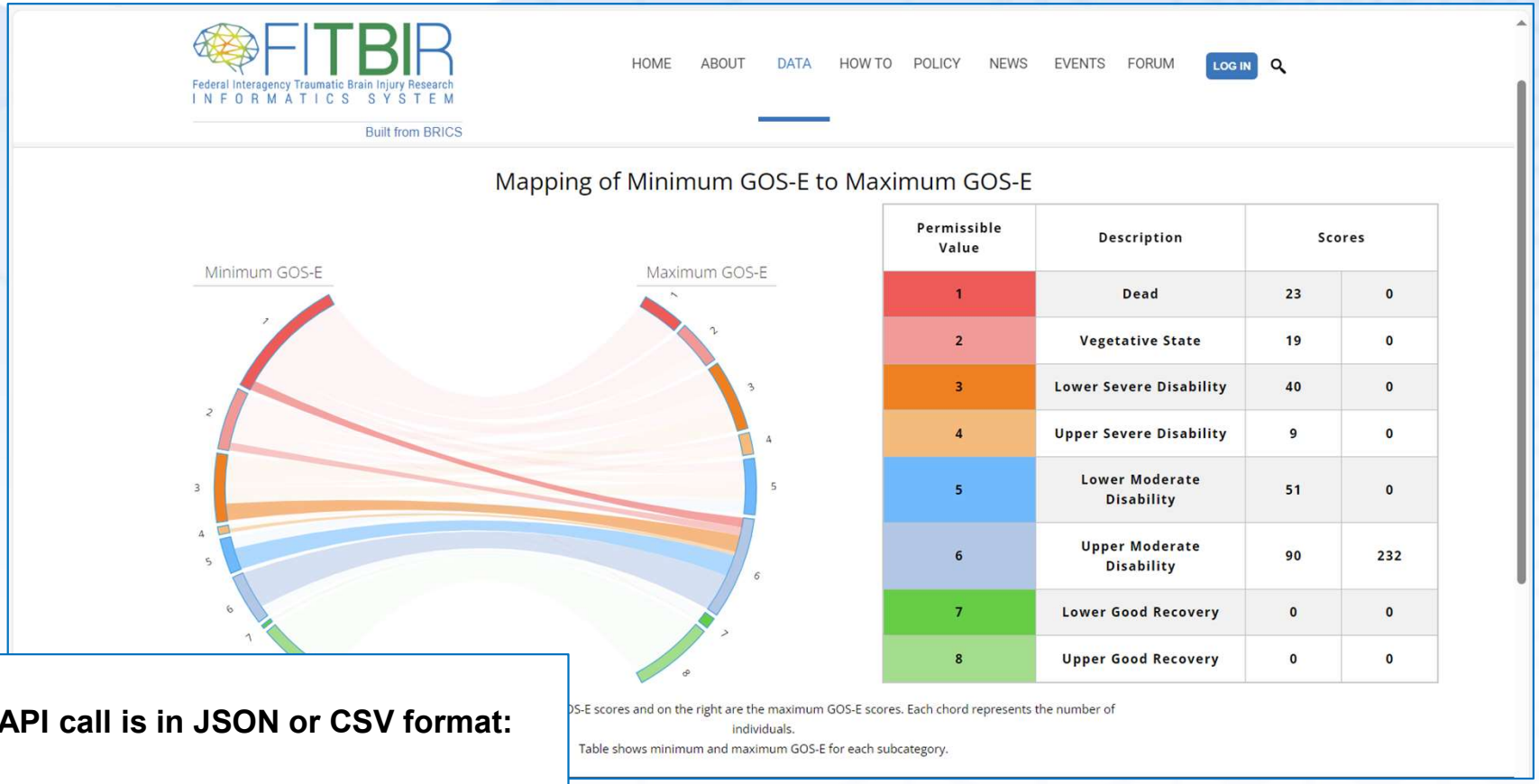
BRICS | Biomedical Research Informatics Computing System

# What can we do with the API?




Data elements across BRICS sorted by
UMLS Semantic Type
Highlight: **Clinical Attribute**

# How do we use the API?




| Permissible Value | Description | Scores | |
|---|---|---|---|
| 1 | Dead | 23 | 0 |
| 2 | Vegetative State | 19 | 0 |
| 3 | Lower Severe Disability | 40 | 0 |
| 4 | Upper Severe Disability | 9 | 0 |
| 5 | Lower Moderate Disability | 51 | 0 |
| 6 | Upper Moderate Disability | 90 | 232 |
| 7 | Lower Good Recovery | 0 | 0 |
| 8 | Upper Good Recovery | 0 | 0 |

**Output of API call is in JSON or CSV format:**


```
"studyId": "string",
  "forms": [
    {
      "id": 0,
      "shortName": "string",
      "title": "string",
      "version": "string"
    }
  ]
```


New ways of looking at and understanding data: **GOS-E**

# Demo

BRICS | Biomedical Research Informatics Computing System

# Questions?

BRICS | Biomedical Research Informatics Computing System

CiSTAR

cdRNS
Common Data Repository for Nursing Science

National Ophthalmic Disease Genotyping Network
eyeGENE®

MTBI²
Uniformed Services University

PDBP
Parkinson's Disease Biomarkers Program

National Institute of Neurological Disorders and Stroke

FITBIR
Federal Interagency Traumatic Brain Injury Research
INFORMATICS SYSTEM

## Biomedical Research Informatics Computing System (BRICS)

# BRICS APIs and Customizing Analysis Pipelines

Dr. Maria Bagonis
Data Scholar

BRICS | Biomedical Research Informatics Computing System

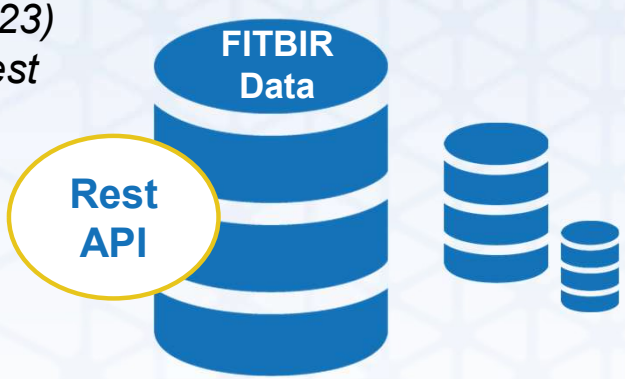
# API Query Tool
# Programmatic access to the data
# For Web Visualization Tools

**BRICS Web Developer**
**Or**
**Researcher**
**Data Scientist**

**Jupyter Notebooks**
**Python, R,**
**Java Script**
**AI/ML libraries**

*Authenticate RAS (11/2023)*
*Submit Query API Request*

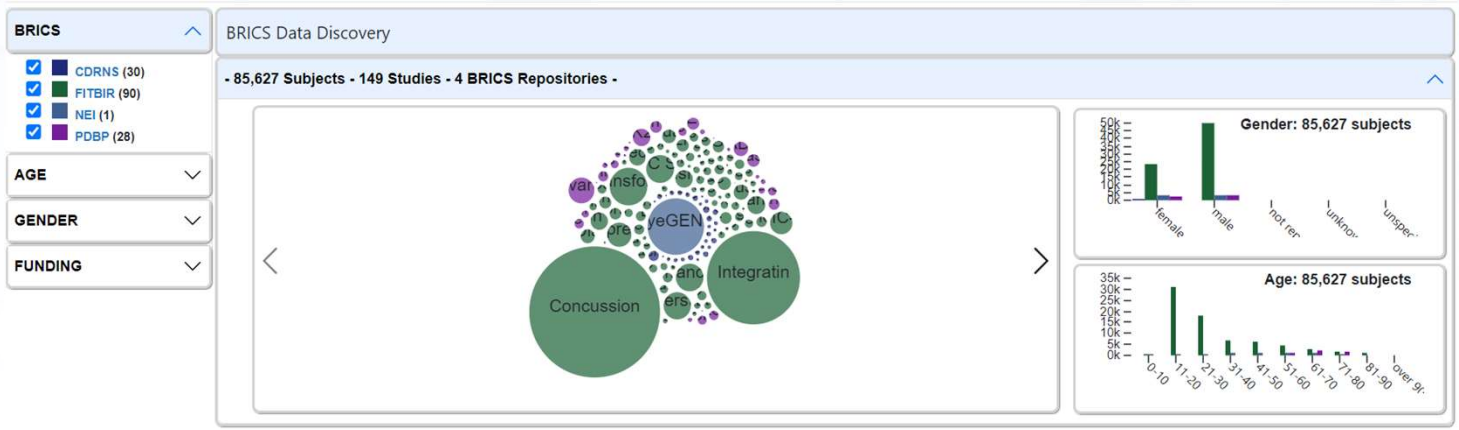
**BRICS Informatics Systems**




*BRICS DATA*

*Retrieve Requested Data*
*Based on Input Parameters*

**Data Discovery Visualization Tools**




https://brics.cit.nih.gov/visualization

BRICS | Biomedical Research Informatics Computing System

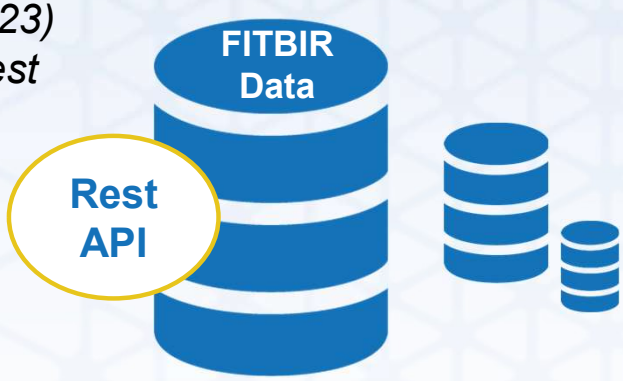
# API Query Tool
# Programmatic access to the data
# For Customized Local Analyses

**Researcher**
**Data Scientist**
**YOU!**

**Jupyter Notebooks**
**Python, R,**
**Java Script**
**AI/ML libraries**

*Authenticate RAS (11/2023)*
*Submit Query API Request*

**BRICS Informatics Systems**

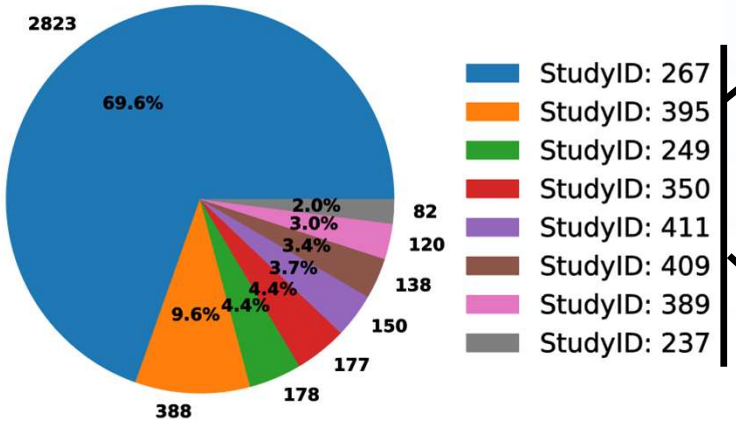


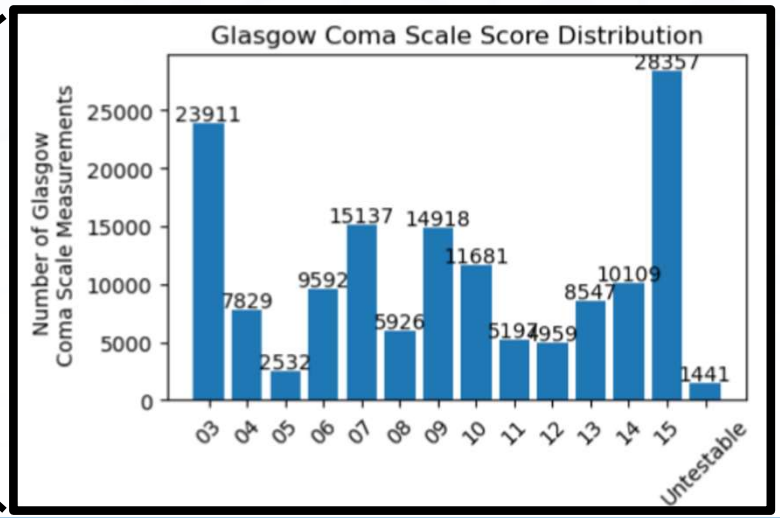

Locally Design Customized Visualizations, Analyses, and Cohort Generation Tools

*BRICS DATA*

*Retrieve Requested Data Based on Input Parameters*

## # of Subjects Per Study Using the GCS Form

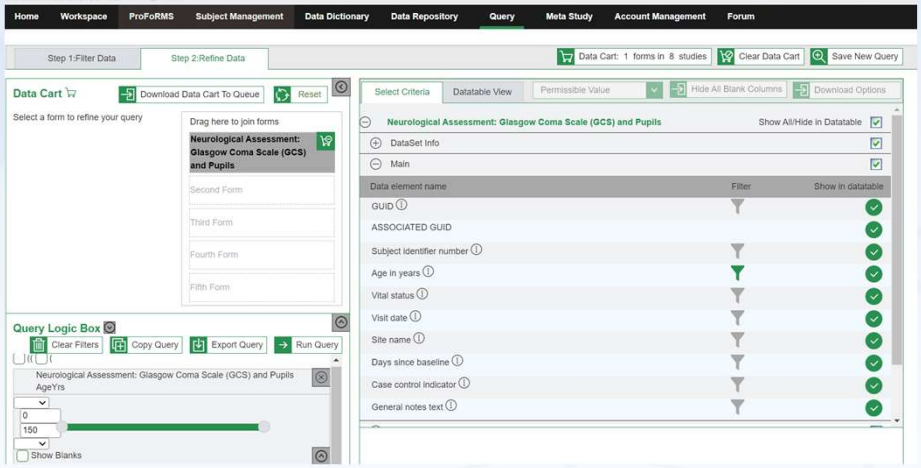


BRICS | Biomedical Research Informatics Computing System

# BRICS Query Tools

## Browser-Based Query Tool




**User Guide**: https://brics.cit.nih.gov/sites/brics/files/2024-02/query-and-api-to-data-query-user-guide.pdf

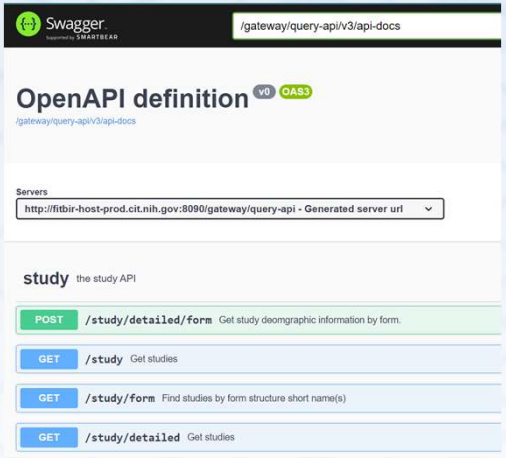
**Tutorial Videos:**

**BRICS Query Tool Introduction:**
**https://www.youtube.com/watch?v=pDukXHWfIBw**

**BRICS Query Tool Step1: Filter Data:**
**https://www.youtube.com/watch?v=uJ0HeLEGVQA**

**BRICS Query Tool Step2: Refine Data:**
**https://www.youtube.com/watch?v=8sL45VF7bx0**

## API Endpoints




**User Guide**: https://brics.cit.nih.gov/sites/brics/files/2024-02/query-and-api-to-data-brics-api-user-guide.pdf

**API Swagger Definitions: (ie where to find API endpoints)**

**FITBIR**: https://**fitbir.nih.gov**/gateway/query-api/swagger-ui/index.html

**PDBP:** https://**pdbp.ninds.nih.gov**/gateway/query-api/swagger-ui/index.html

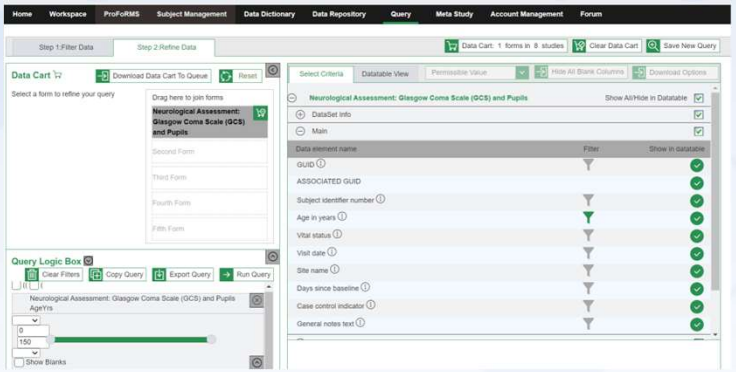
**NEI**: https://**brics.nei.nih.gov**/gateway/query-api/swagger-ui/index.html

**CdRns:** https://**cdrns.nih.gov**/gateway/query-api/swagger-ui/index.html

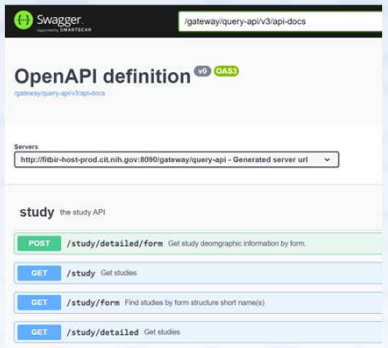
BRICS | Biomedical Research Informatics Computing System

# BRICS Query Tools

## Browser-Based Query Tool




## API Endpoints




## Designed with similar functionality

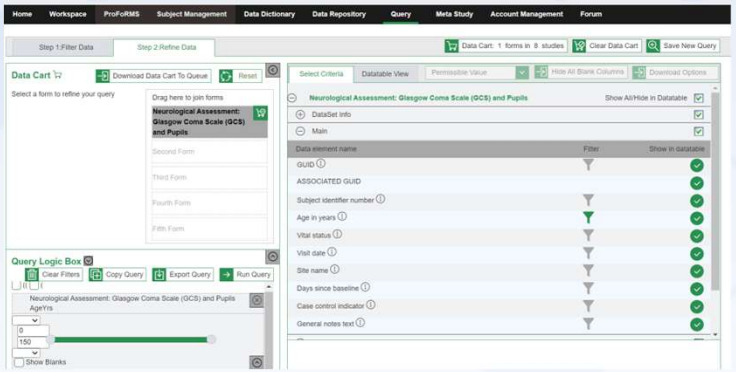
Things you can do:

- Find all form structures submitted for a study
- Find all studies using a specific form structure *(ex. Glasgow Coma Scale (GCS) form in FITBIR)*
- Query data across multiple studies and form structures
  - Filter results based on data element values *(ex. Age, Days Since Baseline, GCSTotal Score*)
  - Perform joins across form structures *(ex. GCS and Diffusion Imaging in FITBIR)*
- Download files associated with query results **Note: We are actively working on workflows for numerous large files (ie, imaging, etc)*

BRICS | Biomedical Research Informatics Computing System

# Difference between Browser and API Query Tool

## Browser-Based Query Tool



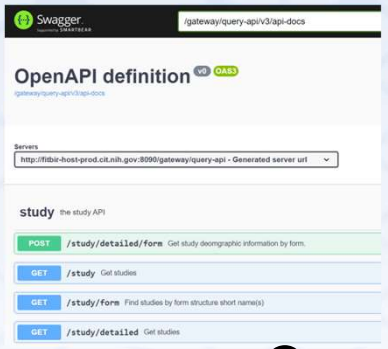

Graphical Interface Query

Check out BRICs Data to Cart

Download BRICs Data via Browser

Upload BRICs data into analysis software (Jupyter Notebook etc)

## API Endpoints




*Allows Movement toward Automated Workflows/Pipelines

Command Line Query from analysis software (Jupyter Notebook etc)

Load BRICS data directly into analysis software (Jupyter Notebook etc)

# Why would we want to use the API?

- **The API provides programmatic access to the Query Tool**
  - Data accessors can use the Query Tool services in their own programs & notebooks

- **Using the API extends the power of the Query Tool**
  - **Directly download** query results to any analysis software that supports REST APIs (Jupyter Notebook/python, R, java-script, etc)
  - **Instantly create** plots of BRICs data returned
  - **Customize** data analysis/transformation pipelines and easily apply them to multiple queries
  - **Share** archived queries and analyses with others.

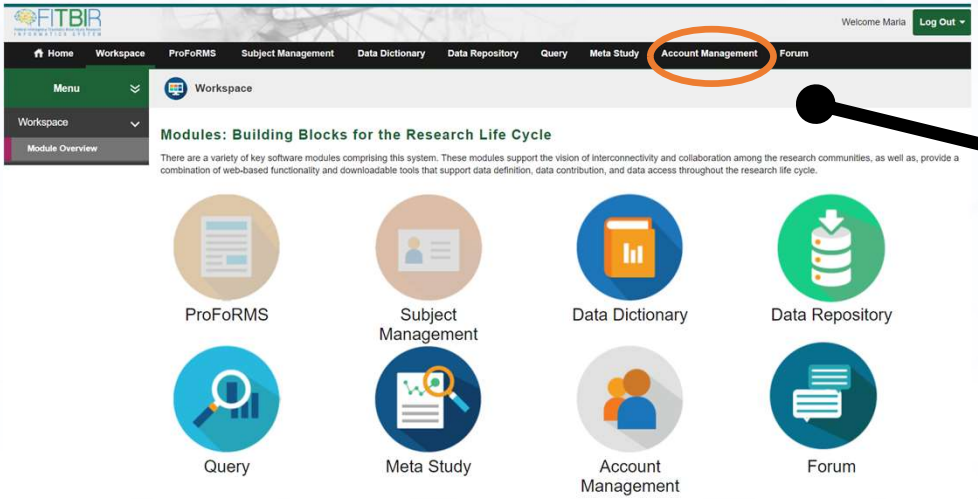
BRICS | Biomedical Research Informatics Computing System

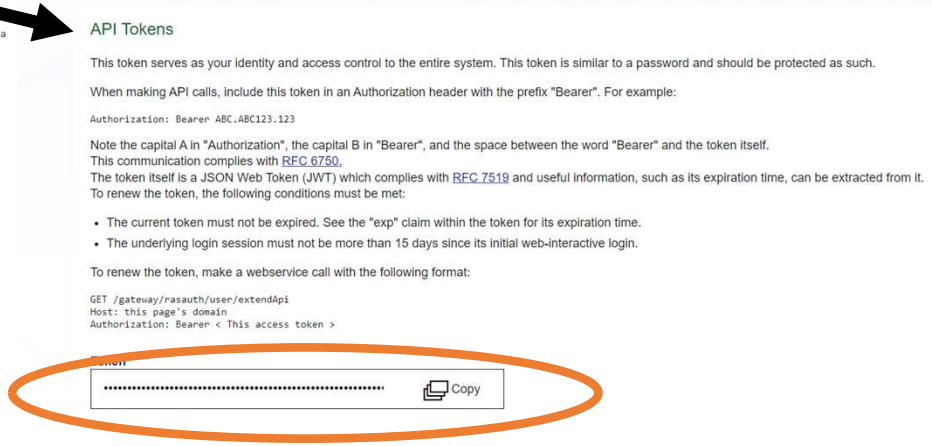
# Steps to Using the API: Step I

## Step I: Define Query URL (ie 'API endpoint'), Header, and Authorization Token

```
# Define API headers and query URL

queryurl ="https://fitbir.nih.gov/gateway/query-api/data/csv"

#Typically Included in Request Header
#Accept - For specifying the type of content that can be accepted in a response. (json etc) See cor
#Content-type - For indicating the type of content that's included in the request body. (json etc)
#Authorization - in the format of 'Bearer ' + token

headers = {
    'accept': 'application/zip', # note this cannot be 'application/csv' as it was in the old examp
    'Content-type': 'application/json',
    'Authorization':'Bearer ' + token }
```

https://fitbir.nih.gov/gateway/query-api/swagger-ui/index.html#/data/getInstancedDataCsv



# Steps to Using the API: Step II and III

## Step II: Define the Request Body (i.e. the Query Filter Information)

```
#
fitbirFormFilter = {
    "formStudy": [
        {
            "form": formName
        }
    ]
}
```

**Defined on Swagger Website:**

Ex. See Here https://fitbir.nih.gov/gateway/query-api/swagger-ui/index.html#/data/getInstancedDataCsv for more information regarding acceptable request body formats

## Step III: Run the Query

```
#Post filter to Query API-
%time query = requests.post(queryurl,headers=headers,json=fitbirFormFilter)
```

BRICS Data →

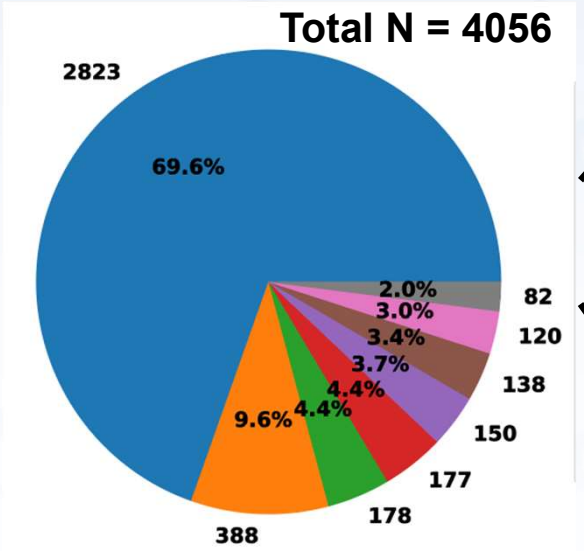
**Develop Local Custom Visualization, Analyses, Cohort Generation Pipelines!**

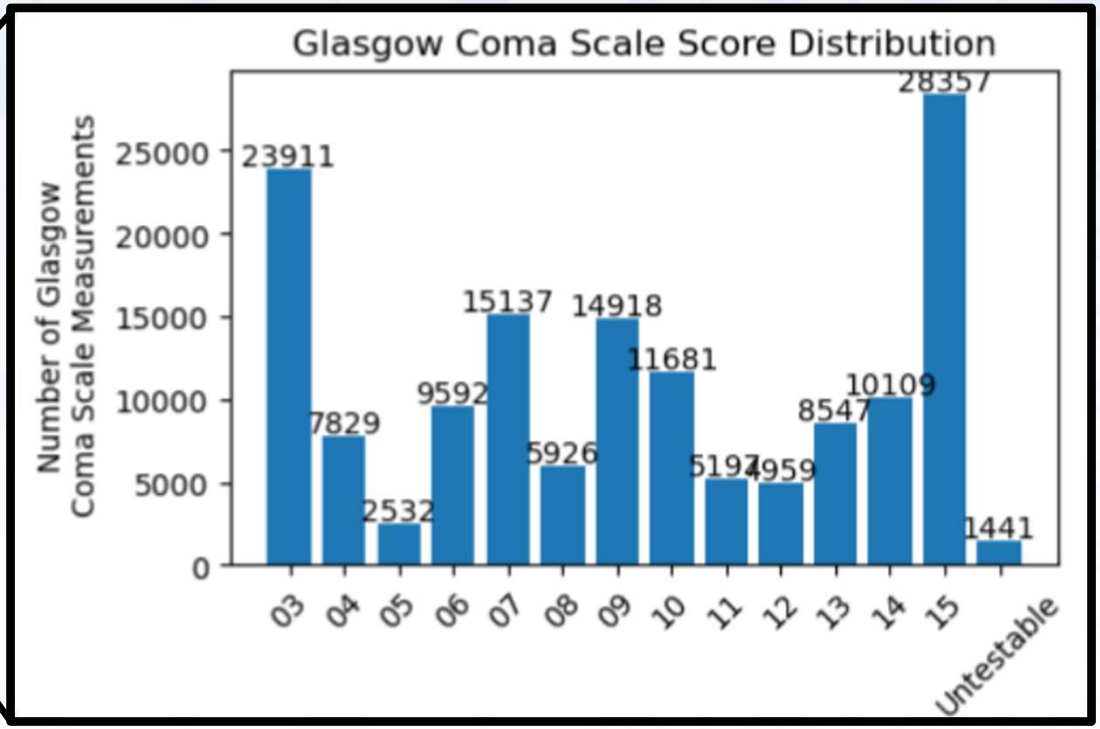
BRICS | Biomedical Research Informatics Computing System

# API Query Tool Examples of Customized Visualization

Find all studies in an instance of BRICs (*ex. FITBIR*) using a specific form *(ex. Glasgow Coma Scale)*
View the Distribution of Data Element Values *(ex.GCS Total Score)* among those studies

## # Subjects Per Study Using GCS Form




## FITBIR Studies

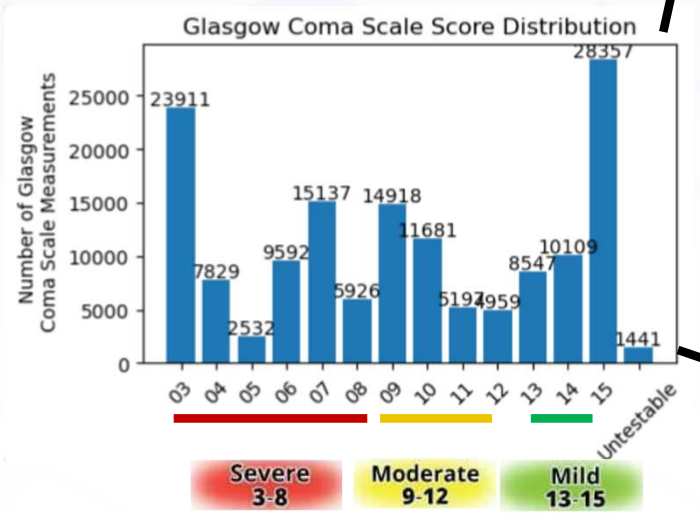
Geoffrey Manley, MD, PhD:Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Adult -StudyID:267
Tellen Bennett:PEDiatric vALidation oF vAriableS in TBI (PEDALFAST) -StudyID:395
Claudia Robertson, MD:Effects of Erythropoietin on Cerebral Vascular Dysfunction and Anemia in Traumatic Brain Injury -StudyID:249
Claudia Robertson, MD:Mission Connect Mild TBI Research Consortium -StudyID:350
David Okonkwo, MD, PhD:Transforming Research and Clinical Knowledge in TBI (TRACK-TBI) - High Definition Fiber Tracking Neuroimaging, Biospecimen, and Data Informatics Repositories -StudyID:411
Geoffrey Manley, MD, PhD:Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Pediatric -StudyID:409
Jed Hartings:Development and validation of spreading depolarization monitoring for TBI management -StudyID:389
Geoffrey Wright:Development and Validation of the Virtual Environment TBI Screen (VETS) for Postural Control Testing -StudyID:237

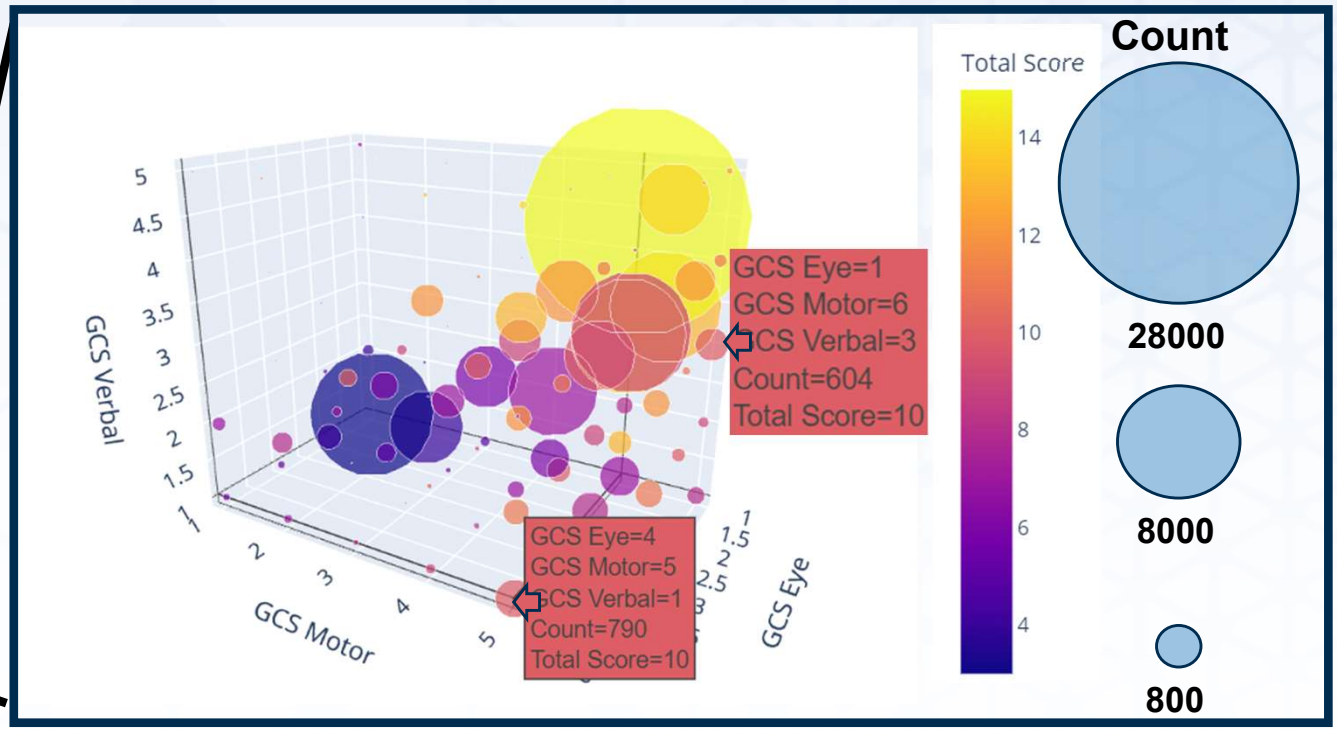
See Python Notebook:
BRICSAPIDemo_CollectDataByForm.ipynb

BRICS | Biomedical Research Informatics Computing System

# API Query Tool Examples of Customized Visualization

Large Sample Sizes Afforded by Combining Data Among Multiple Studies Can Reveal Important Heterogeneity in the Patient Population and Ensure Sampling of Less Common Sub-populations

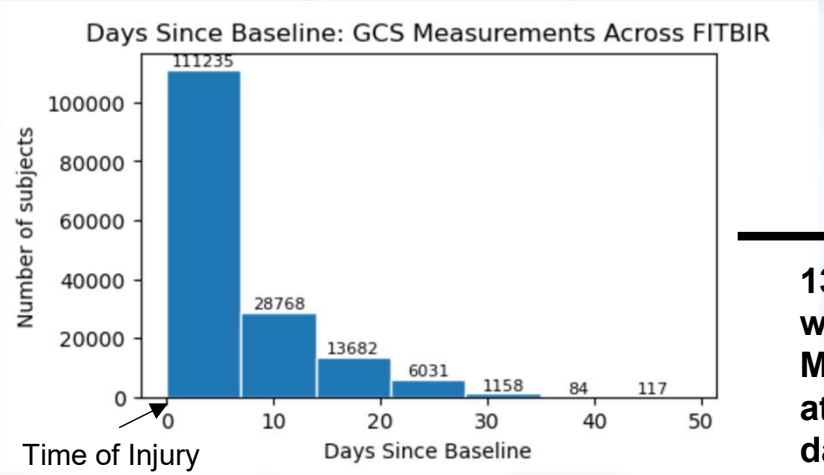

See Python Notebook: BRICSAPIDemo_CollectDataByForm.ipynb

# API Demo

BRICS | Biomedical Research Informatics Computing System

# API Query Tool Examples of Customized Visualization (Extra)

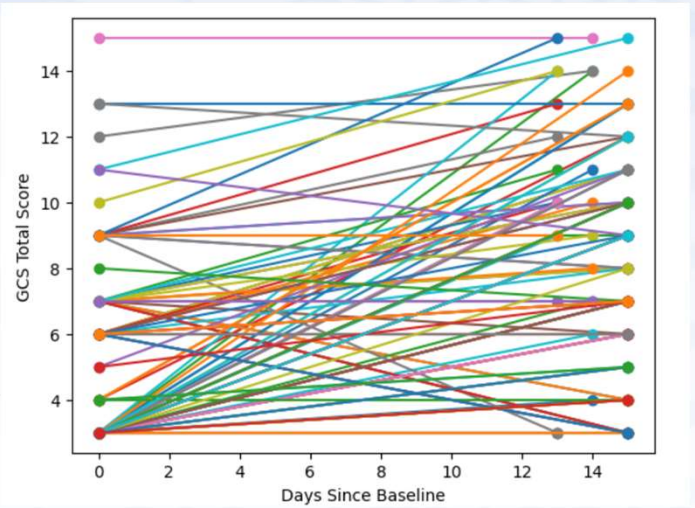
Extracting Temporal Measurements:
Glasgow Coma Scale (GCS) Measurements Taken at Different Times Since Injury



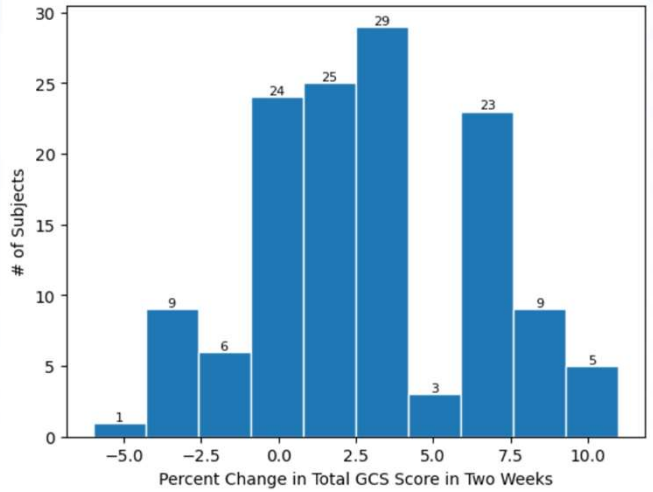

**134 Subjects with GCS Measured BOTH at 0 and ~14 days after injury**

**At Time of Injury**

**Two Weeks Post Injury**




**Worsened** **Improved**

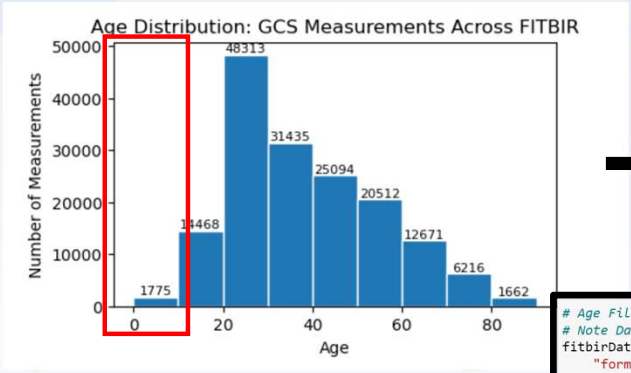



See Python Notebook:
BRICSAPIDemo_CollectDataByForm.ipynb

BRICS | Biomedical Research Informatics Computing System

# API Query Tool Examples of Customized Visualization (Extra)

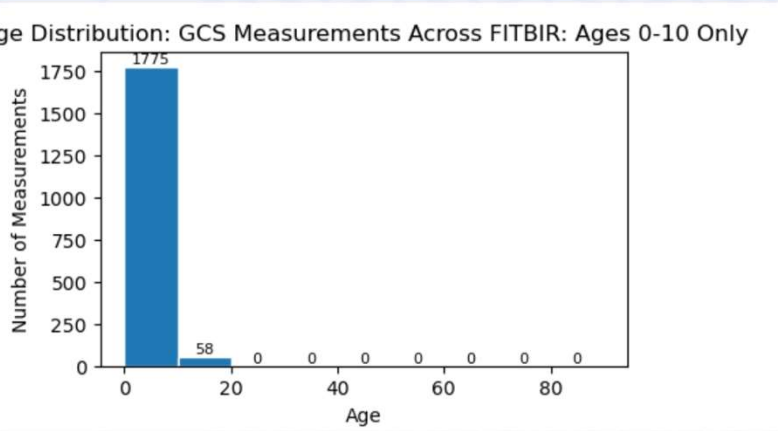
Find all studies using a specific form structure and filter by Data Element Value as defined by form.repeatableGroup.dataElement (i.e GCS.Main.AgeYrs)



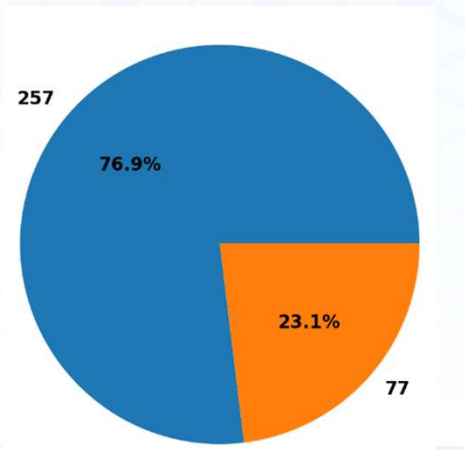

Create Pediatric Filter

Only Download Data: Ages 0-10

```
# Age Filter 0-10
# Note Data Element Defined  formName.repeatableGroup.dataElementName: Ex GCS.Main.AgeYrs
fitbirDataElementFilter = {
    "formStudy": [
        {
        "form": GCS #"string",
#           "studies": [
#             "string"
#           ]
        }
        ],
       "filter": [
         {
          "form": 'GCS', # string
          "repeatableGroup": 'Main' , #"string"
          "dataElement": 'AgeYrs', #"string"
          "rangeStart": "0", #string
          "rangeEnd": "10", #string
          "mode": "inclusive",
         }
       ],
}
```




## # Subjects Per Study Using GCS Form




Study

Tellen Bennett:PEDiatric vALidation oF vAriableS in TBI (PEDALFAST) -StudyID:395

Geoffrey Manley, MD, PhD:Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Pediatric -StudyID:409

See Python Notebook:
BRICSAPIDemo_CollectDataByForm_WithDataElementValueFilter.ipynb

# API Query Tool Examples of Customized Visualization (Extra)

## Find all studies using BOTH ImagingDiffusion and GCS forms:

| id | title | pi |
|---|---|---|
| FITBIR-STUDY0000411 | Transforming Research and Clinical Knowledge in TBI (TRACK-TBI) - High Definition Fiber Tracking Neuroimaging, Biospecimen, and Data Informatics Repositories | David Okonkwo, MD, PhD |
| FITBIR-STUDY0000409 | Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Pediatric | Geoffrey Manley, MD, PhD |
| FITBIR-STUDY0000267 | Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Adult | Geoffrey Manley, MD, PhD |

## Find all studies using BOTH Biomarker and ImagingFunctionalMR forms:

| id | title | pi |
|---|---|---|
| FITBIR-STUDY0000411 | Transforming Research and Clinical Knowledge in TBI (TRACK-TBI) - High Definition Fiber Tracking Neuroimaging, Biospecimen, and Data Informatics Repositories | David Okonkwo, MD, PhD |
| FITBIR-STUDY0000310 | Concussion Assessment, Research and Education (CARE) Consortium | Steven P Broglio |
| FITBIR-STUDY0000409 | Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Pediatric | Geoffrey Manley, MD, PhD |
| FITBIR-STUDY0000279 | Comprehensive study of acute effects and recovery after concussion: An integrated investigation of head impact sensor technology, blood biomarkers, advanced neuroimaging, genetic testing and clinical outcome metrics | Michael McCrea |
| FITBIR-STUDY0000267 | Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) - Adult | Geoffrey Manley, MD, PhD |
| FITBIR-STUDY0000384 | TBI Endpoints Development (TED) | Geoffrey Manley |

See Python Notebook: BRICSAPIDemo_FindStudiesUsingTwoOrMoreCommonForms.ipynb

*Not the form you are looking for?: Interactively Search All FITBIR Forms via the web interface here:* https://fitbir.nih.gov/content/data-dictionary#form-structures

# API Query Tool Examples of Customized Visualization

## Insert Your Visualization Here!

**BRICS API Manual (PDF):** https://brics.cit.nih.gov/sites/brics/files/2024-02/query-and-api-to-data-brics-api-user-guide.pdf

**BRICS API Manual (Jupyter Notebook):** https://brics.cit.nih.gov/sites/brics/files/2024-02/query-and-api-to-data-brics-api-user-guide.ipynb

**All demo notebooks from today's webinar are located at** https://brics.cit.nih.gov/documentation in the **Query and API to Data - API Scripts Zip File**

*Questions/Comments?*

Additional Questions/Comments/Suggestions
Contact: maria.bagonis@nih.gov
*Share your voice and be part of the community!*

BRICS | Biomedical Research Informatics Computing System

# Closing Remarks