

# Real-Time Integration of EDC and Third-Party Data and Query Actions within an R Shiny Data Review and Patient Profile Tool

March 2026



# Agenda

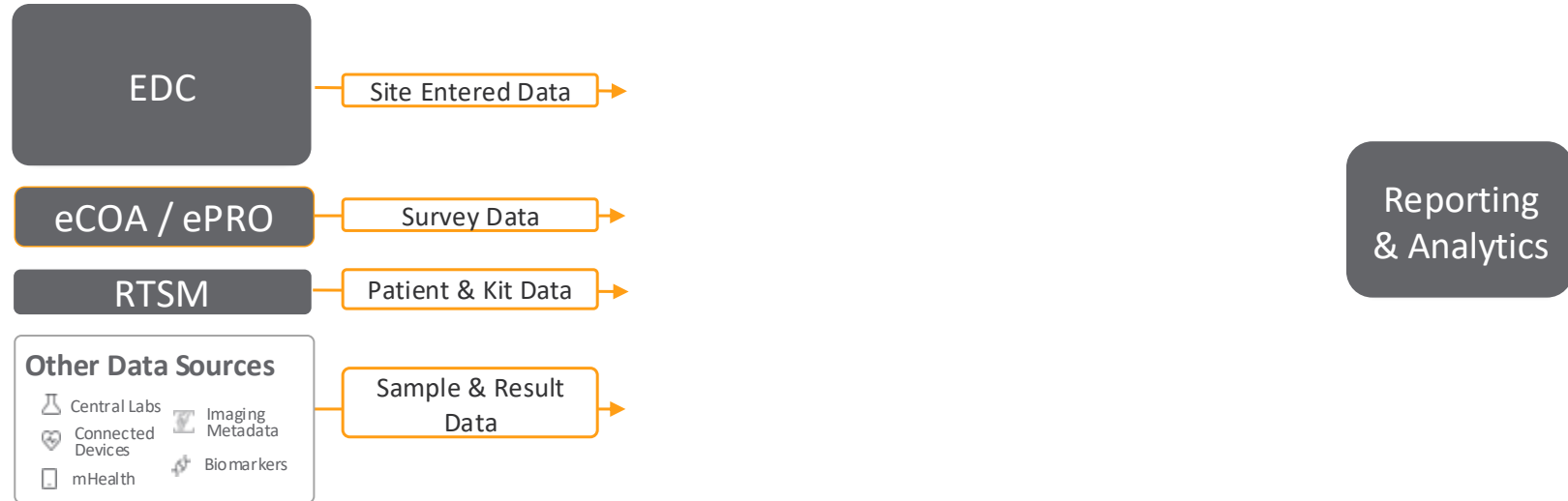
- Why legacy data review methods are no longer good enough
- How modern incremental data APIs and API driven actions meet today's needs
- Architecting an open source data review tool with integrated actions
- Demonstration



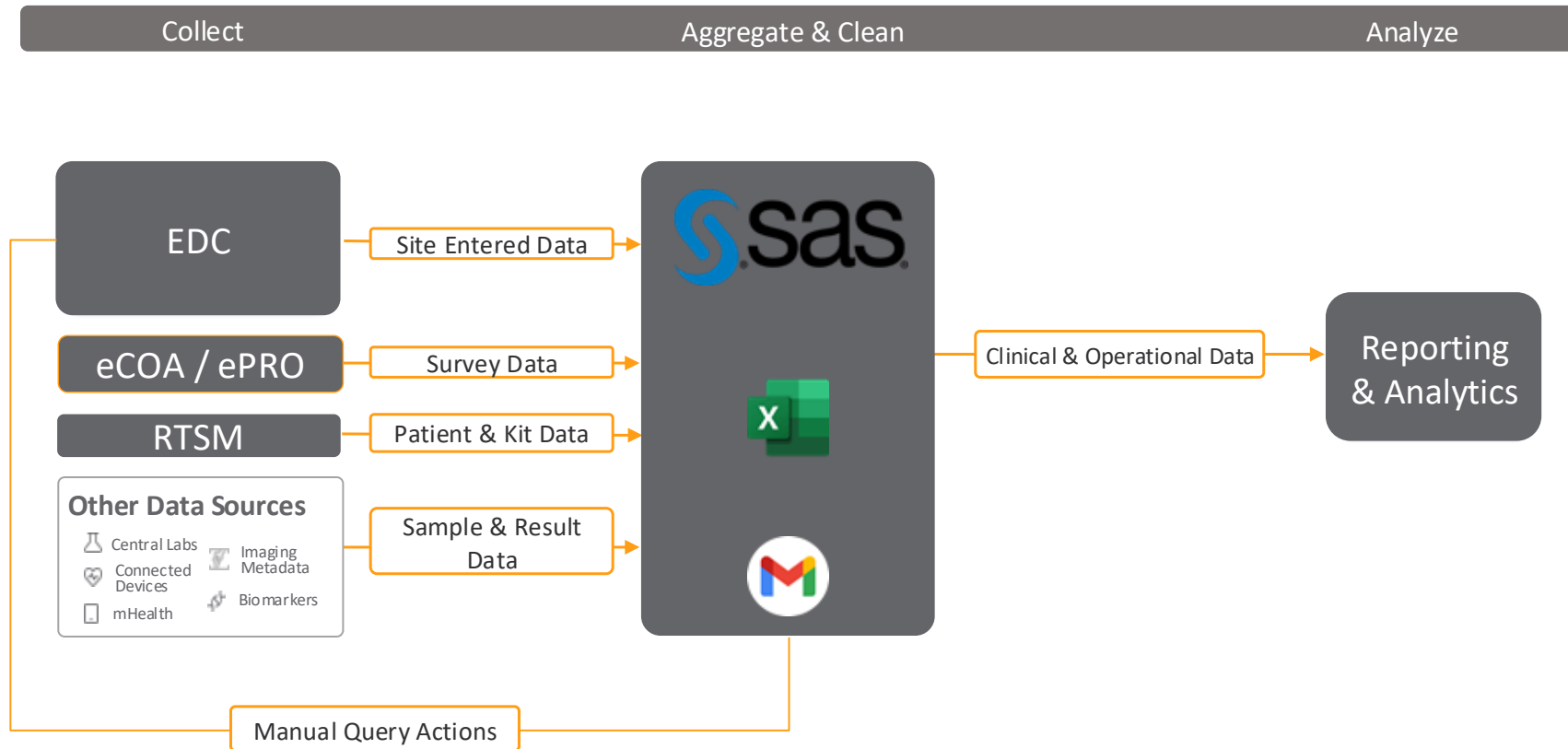
# Clinical trial teams need rapid access to data from multiple systems

Collect

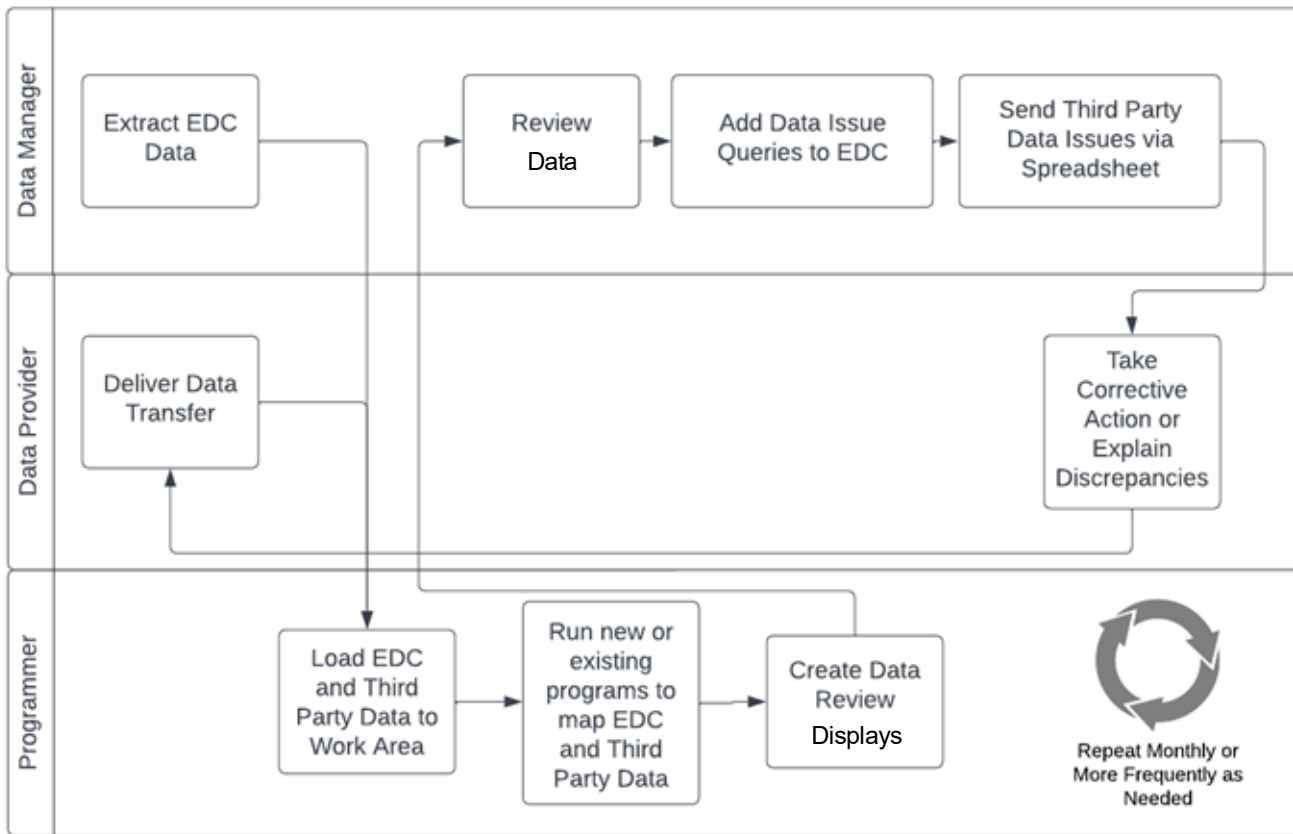
Analyze



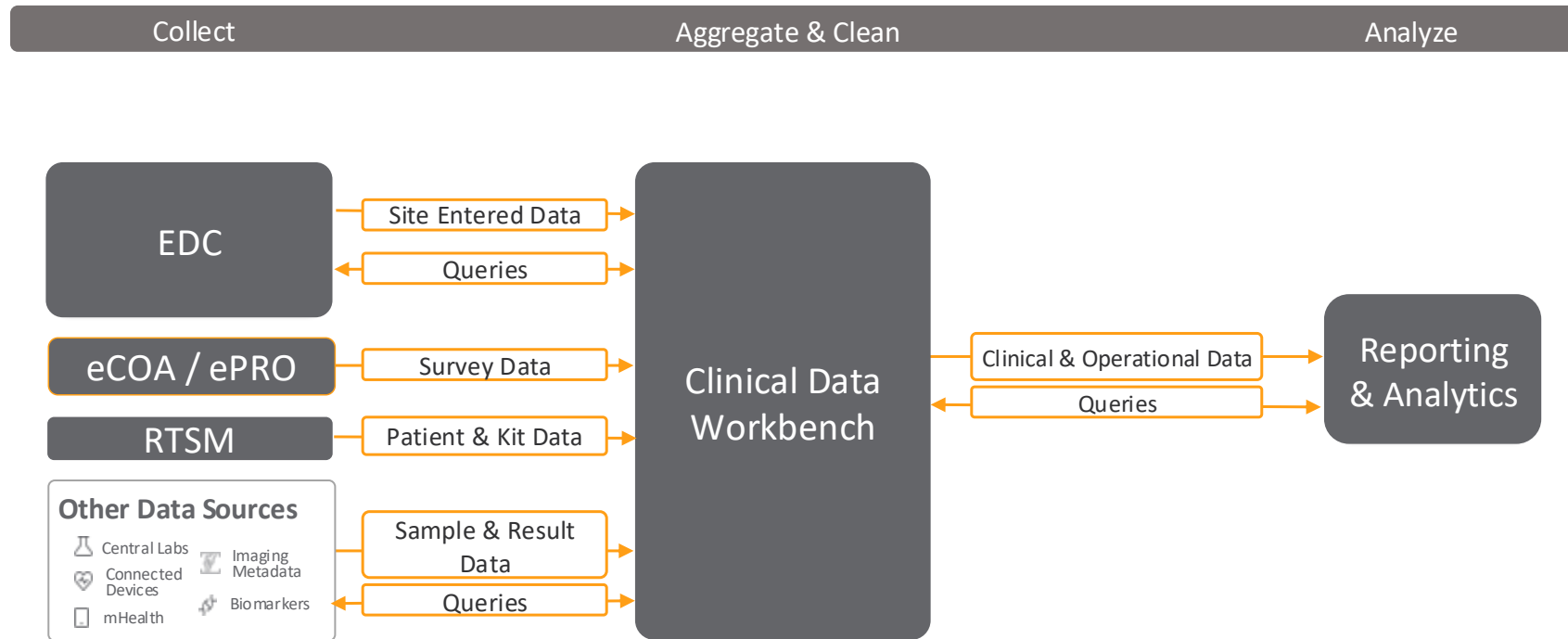
# But first that data needs to be taken offline, aggregated, and cleaned using non-GxP compliant tools



# Legacy data review is manual and error-prone



# Clinical data workbenches and open source reporting and analytics tools can support a GxP compliant data review workflow



# Veeva Study File Format API is used for live, real-time integration

## What is it:

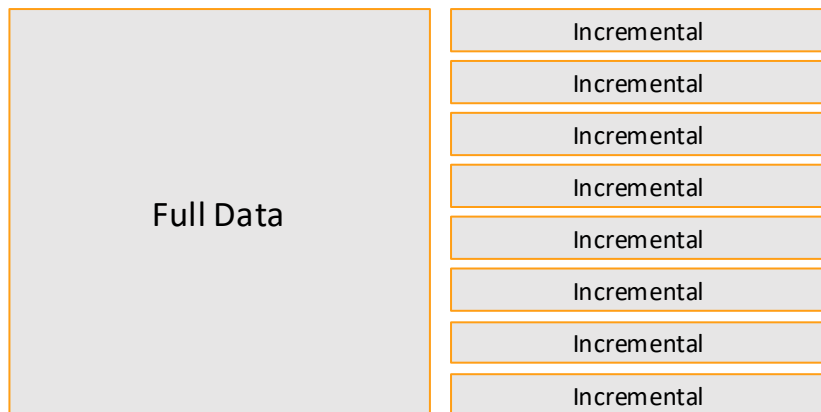
- A full and incremental export of the study's **clinical, operational, and reference data**
- Self-describing payload is built for metadata-driven programming
- Includes study design metadata in the manifest.json file
- Available via REST API

## When to use it

- When you want rapid and incremental data synchronization to an external system
- When you want a completely programmatic export solution
- When you want a standardized export format out of EDC & CDB for your clinical, operational, and reference data
- When you want a single export package with EDC and non-EDC data ingested into CDB (25R1)



# Working with Incremental Data



- Normal Application
  - Start with a full
  - Apply changes from incrementals every 15 mins
  - Delete records flagged for delete.
- Design Change
  - Wait for next full
  - Update Schema with changes
  - Apply Full
  - Apply Incrementals
- Too Fast?
  - Every hour grab the 4 incrementals.
  - Merge Incrementals
  - Dedupe Incrementals
  - Apply Larger Incremental

# Metadata-driven approaches are enabled by a manifest.json file

```
{
  "study" : "Natevba_Data",
  "study_external_id" : "Natevba_Data",
  "documentation" :
    "https://cdmshelp.veeva.com/gr/monitor
    s/cdb/study-file-format/",
  "version" : 1.0,
  "extract_name" :
    "Natevba_Data_SFF_Full_2025_05_22_12_0
    0_00.zip",
  "extract_created_date" : "2025-05-
    22T06:00Z",
  "file_count" : 72,
  "incremental" : false,
  "study_design_version" : 2,
  "clinical_data" : [ {
    "filename" : "AEYN.csv",
    "source" : "EDC",
    "form" : "AEYN",
    "local_lab_form" : false,
```

Full file / payload  
definition

# Including deep column-level definitions

```
"columns" :  
[ {  "name" : "STUDYNAME",  
    "datatype" : "text",  
    "precision" : null  },  
  {  "name" : "FIRSTSUBMITDT",  
    "datatype" : "datetime",  
    "precision" : null  },  
  {  "name" : "AEYN",  
    "datatype" : "codelist",  
    "precision" : null,  
    "itemgroups" : [ "IG-AEYN" ],  
    "external_id" : "AEYN",  
    "sdtn_name" : null,  
    "length" : 1500,  
    "mask" : null,  
    "unknown" : null,  
    "item_type" : [ "edc__v" ],  
    "restricted" : false,  
    "codelist" : "cl-NY",  
    "unit" : null  
  }  
]
```

Column-level  
metadata

# And value-level definitions

## Codelists

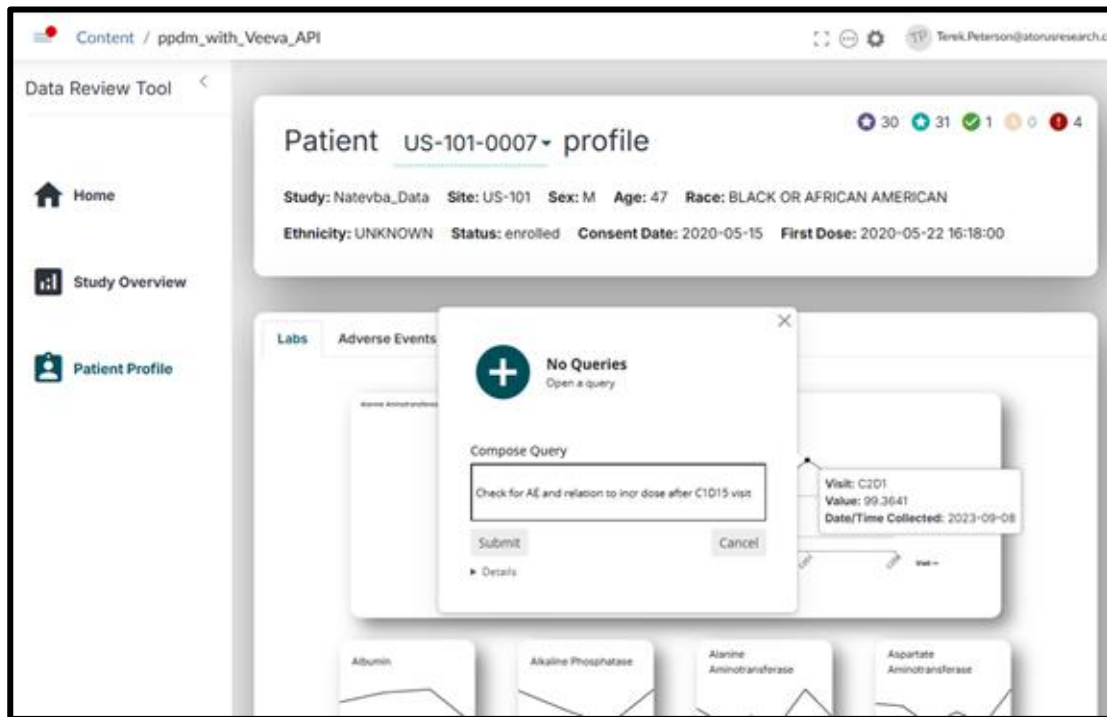
```
{ "name" : "cl-NY",  
  "external_id" : "cl.NY",  
  "source" : "EDC",  
  "codelist_data" : [ {  
    "code" : "Y",  
    "decode" : "Yes"  
  }, {  
    "code" : "N",  
    "decode" : "No"  
  } ]  
}
```

## Units and conversions

```
"units" : [ {  
  "name" : "ucHeight",  
  "external_id" : "ucHeight",  
  "source" : "EDC",  
  "type" : "length__v",  
  "unit_data" : [ {  
    "code" : "in",  
    "decode" : "inches",  
    "external_id" : "in",  
    "abbreviation" : null,  
    "standard" : false,  
    "conversion" : "$value__v * 2.54"  
  } ]  
}]
```

# R Shiny data review tool: an interactive data product

- Users authenticate using Posit Connect
- Data security and blinding control
- GxP compliant role management
- Activity logging
- Integration actions via Query API



# Query API delivers queries to site's EDC in real time

Subject: US-101-0001

Sort By: Schedule Sign + New Event

Common Forms: Adverse Events: Reference #1 of 5

FORM | PRIOR AND CONCOMITANT ... (0)

Adverse Events

Adverse Event

hypotension

Start Date

02-Jun-2020

Terek.Peterson@atorusresearch.com (ppdm app)

02 Sep 2025 3:12 PM EDT

AEESTDAT is after AEENDAT

Edit Value Comment

Ongoing?

Yes

Severity

Mild

Mike.Stackhouse@atorusresearch.com (ppdm app)

30 Oct 2025 2:20 PM EDT

This wasn't severe

Edit Value Comment

Terek.Peterson@atorusresearch.com (ppdm app)

11 Sep 2025 8:47 AM EDT

Severity is Mild but Patient was Hospitalized with Serious AE, confirm

Edit Value Comment

COMMON FORMS

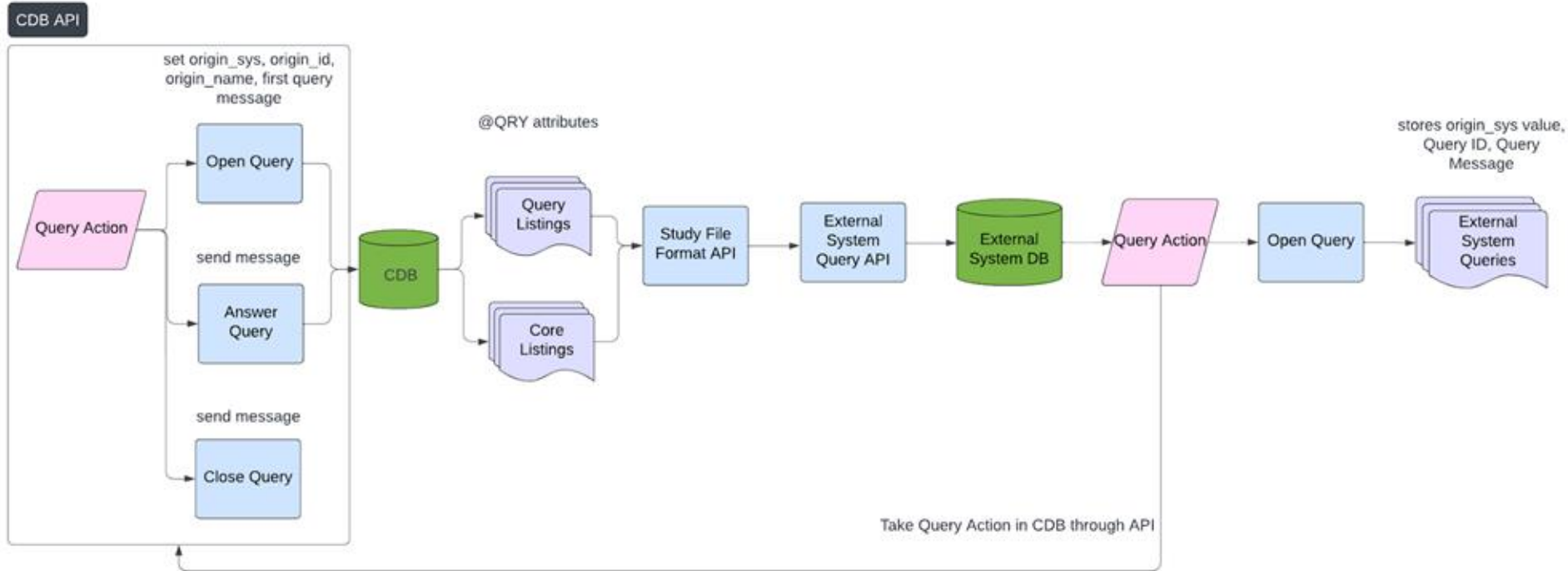
Adverse Events YN?

Prior and Concomitant Medications YN?

Adverse Events (5)

Prior and Concomitant Medications (1)

# Query API Workflow



# Future Enhancements

- Veeva APIs continue to introduce new endpoints for integrated actions
  - Queries (available now)
  - Protocol Deviations (future)
  - Cross-functional observations (future)
- Data review tool will enhance metadata-driven dynamic capabilities
  - Out-of-the-box visualizations based on the class or domain of data
  - Aggregate data analytics, e.g. signal detection
  - Snapshotting of data



# Conclusion

- Legacy data exchange mechanisms are uni-directional and cause data lag
- Modern API-based data exchange methods like SFF API and query API support bi-directional actions like queries and incremental data flow
- The clinical data workbench excels at data aggregation, reconciliation and cleaning
- The R Shiny data review tool provides needed flexibility and customization for different stakeholders to interact and explore their data with visualizations





# Questions?

## **CONTACT INFORMATION**

Author Name: Pavel Burmenko

Company: Veeva Systems

Email: [Pavel.Burmenko@veeva.com](mailto:Pavel.Burmenko@veeva.com)

Author Name: Terek Peterson

Company: Atorus Research

Email: [Terek.Peterson@atorusresearch.com](mailto:Terek.Peterson@atorusresearch.com)

