# RADx-rad Data File Organization and Data Elements

Data File Organization

The data for RADx-rad diagnostic methods development are organized into multiple data files. Not all studies follow this schema and some contain a subset of these data files.

Diagnostic methods have been benchmarked using either clinical or spiked samples, or both, which is represented in this schema by two parallel tracks.
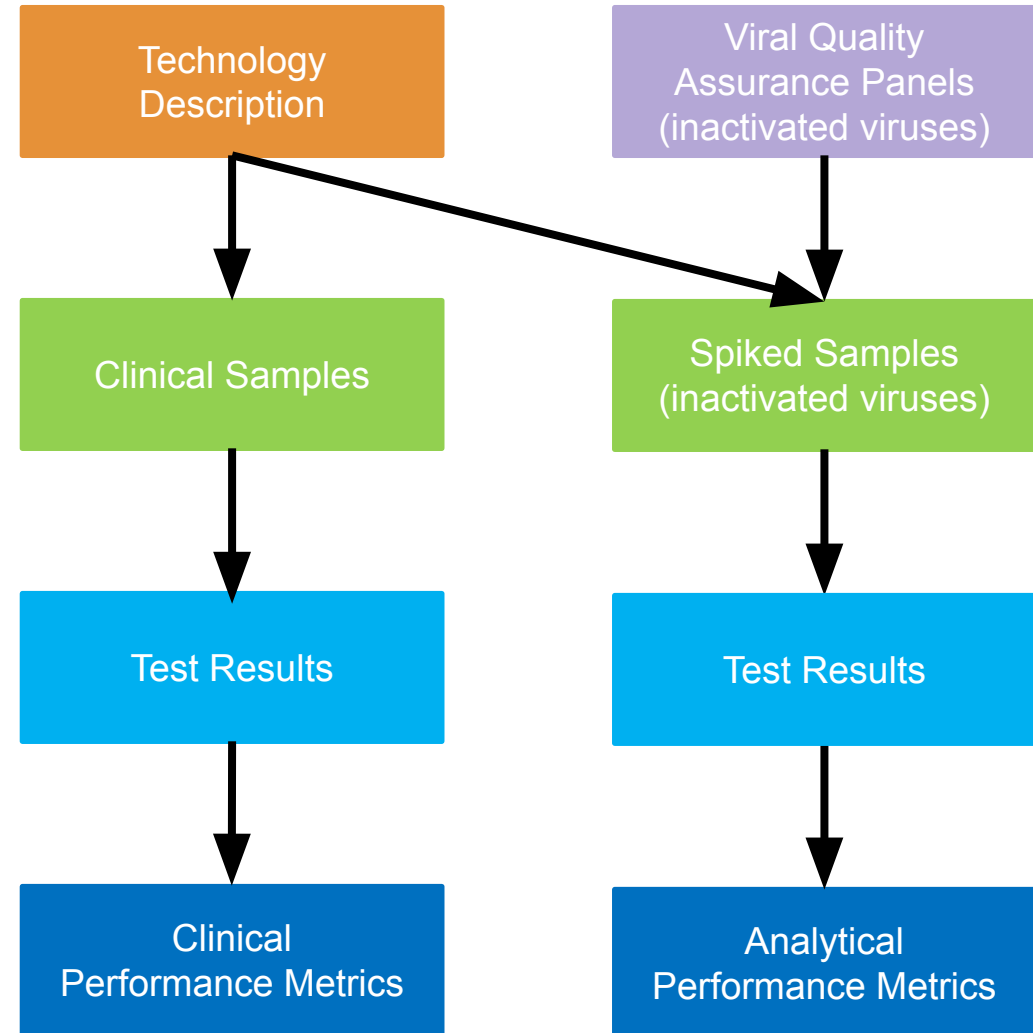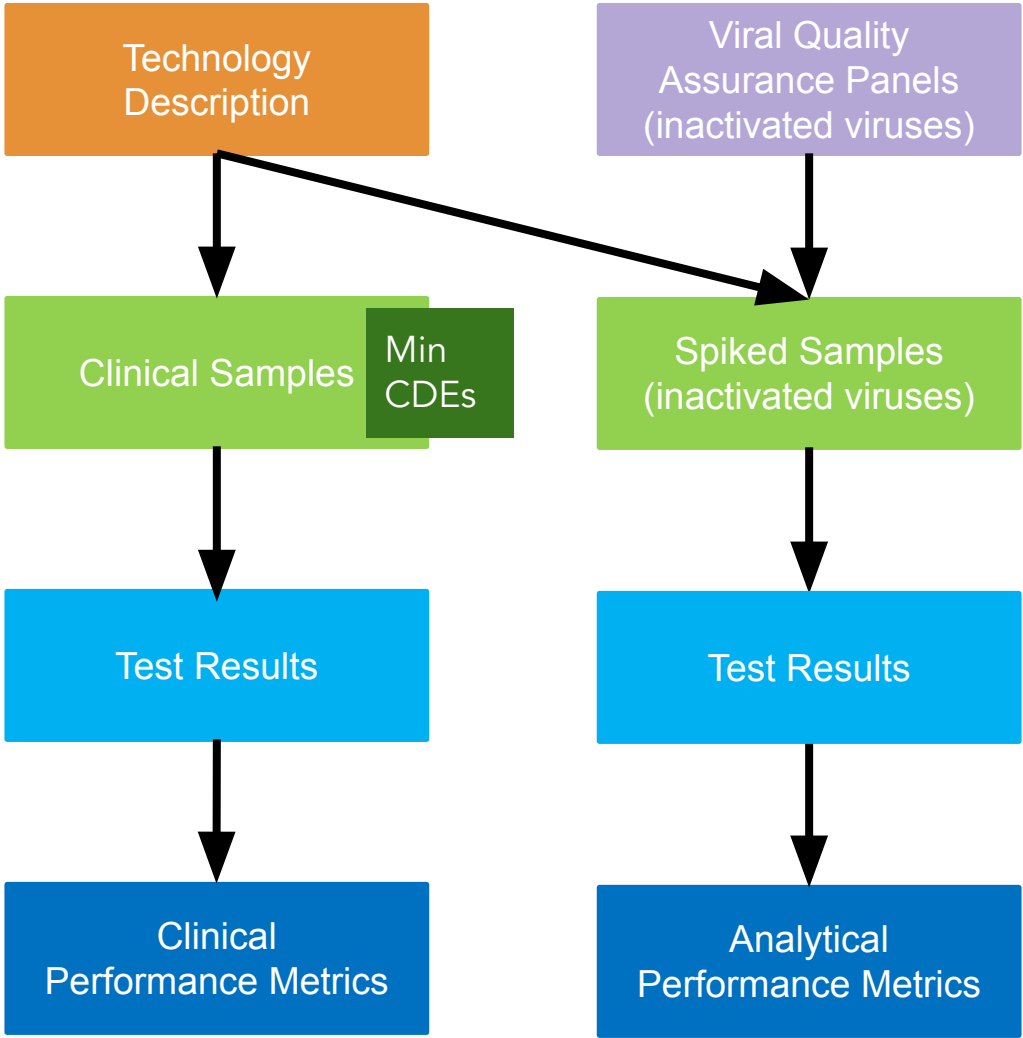
File Content

Methods and materials

Demographics, medical history, symptoms, COVID status, etc. for de-identified subjects. They may be grouped into cohorts.
Min CDEs are a subset of harmonized data elements.

Method or assay readouts for each subject

Performance metrics such as sensitivity and specificity for each cohort

Lineage, source, formulation of inactivated viruses

Concentration and composition of samples. Samples may be grouped into sample groups.

Method or assay readouts for each sample

Performance metrics such as limit of detection for each sample group

Technology Description

Viral Quality Assurance Panels (inactivated viruses)

Clinical Samples   Min CDEs

Spiked Samples (inactivated viruses)

Test Results

Test Results

Clinical Performance Metrics

Analytical Performance Metrics

References among Data Files

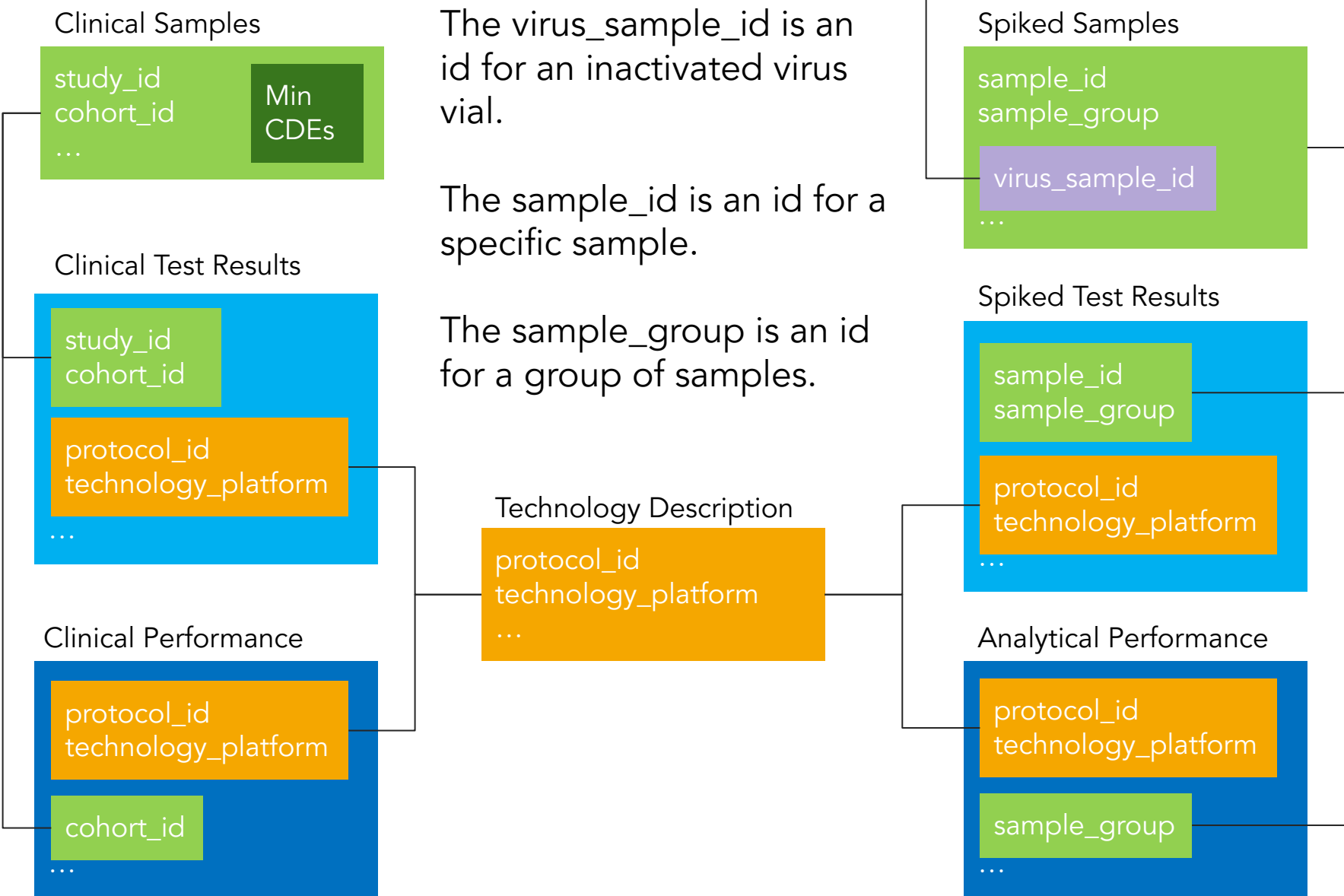Unique identifiers establish references among the data files.

The study_id is a de-identified subject id.

The cohort_id is an id for a cohort of subjects.

The protcol_id is an id for a specific technology protocol.

The technology_platform is a short name for the technology/protocol.

The virus_sample_id is an id for an inactivated virus vial.

The sample_id is an id for a specific sample.

The sample_group is an id for a group of samples.

Clinical Samples
study_id
cohort_id
...
Min CDEs

Clinical Test Results
study_id
cohort_id

protocol_id
technology_platform
...

Clinical Performance
protocol_id
technology_platform

cohort_id
...

Technology Description
protocol_id
technology_platform
...

Virus Quality Assurance Panel
virus_sample_id
...

Spiked Samples
sample_id
sample_group

virus_sample_id
...

Spiked Test Results
sample_id
sample_group

protocol_id
technology_platform
...

Analytical Performance
protocol_id
technology_platform

sample_group
...

References

Tier 1 and Tier 2 Data Elements

- Tier 1 data elements (Min CDEs) are a common harmonized subset of data elements that describe demographics, medical history, symptoms, and health status of subjects
- Tier 2 data elements are used to describe RADx-rad specific data
- For the definition of these data elements see:

https://github.com/radxrad/common-data-elements