# RADX-UP Data Quality Resolution Policy

## I. Introduction and Purpose

The National Institutes of Health (NIH) launched the Rapid Acceleration of Diagnostics (RADx) initiative to speed innovation in the development, commercialization, and implementation of technologies for COVID-19 testing. As part of this initiative, NIH developed the RADx-UP (RADx-Underserved Populations) program to understand disparities in COVID-19 morbidity and mortality and to help reduce disparities for underserved populations who are disproportionately affected by the COVID-19 pandemic.

Data from the RADx-UP studies will be transferred to the RADx-UP Coordination and Data Collection Center (RADX-UP CDCC) and aggregated to help researchers answer broad questions that the individual studies cannot answer alone. The purpose of the RADx-UP Data Quality Resolution Policy is to ensure the datasets submitted to the RADX-UP CDCC conform with program data formatting and quality guidelines.

The following principles guided the development of the policy:

- Sharing data in a timely manner is important for optimizing scientific progress, especially in the time of the COVID-19 pandemic.
- This policy is consistent with the goals of the NIH data-sharing policy to make public access to digital scientific data the standard for all NIH-funded research. See Draft NIH Policy for Data Management and Sharing.
- Combining and harmonizing data from multiple projects will increase power and generalizability of results for addressing important research questions, particularly for populations typically excluded from analyses due to small numbers.

## II. Scope

This policy applies to all RADx-UP projects that submit data to the RADX-UP CDCC.

## III. Definitions

| Item | Description |
|---|---|
| De-identified Data | The term "*de-identified data*" refers to a set of data where *protected health information (PHI)* is removed from data sets. De-identified data will be those for which all 18 designated HIPAA identifiers (as listed at https://privacyruleandresearch.nih.gov/pr_08.asp) have been deleted or altered to protect the privacy of participants. |

| | |
|---|---|
| DTA/DUA | Data Transfer Agreement / Data Use Agreement |
| EIT | Engagement Impact Team. The two person RADX-UP CDCC team assigned to liaise with an individual project. |
| RADx-UP CDCC Informaticist | The RADx-UP CDCC Informaticists are assigned to work directly with an individual project. Projects will be introduced to their informaticist via their EIT. |
| RADx-UP CDCC | The RADx-UP CDCC is responsible for developing the secure environment for accessing RADx-UP consortium data. It is responsible for managing receipt of *RADx-UP Consortium data* from RADx-UP projects, harmonization of *RADx-UP project data*, conducting analysis, and providing access to the *RADx-UP project data* for querying, visualization, and analysis. |
| RADx-UP Consortium | Includes the RADx-UP CDCC, projects funded in the RADx-UP program, pilot project and community grants from the RADx-UP CDCC, principal investigators and key personnel of the referenced projects, NIH program staff, and project scientists. |
| RADx Data Hub | The NIH data repository for the entire RADx Program. Data from all of the RADx programs (RADx-UP, RADx-rad, RADx-ATP and RADx Tech) will be sent to the RADx Data Hub. |
| RADx-UP Data Portal | A secure web portal maintained *within myRADx-UPhome* by the *RADx-UP CDCC* to ingest, host, and control access to data held in the RADx-UP data platform. The *RADx-UP Portal* provides secure access for consortium members with fully executed Data Use Agreements to the harmonized RADx-UP de-identified data set. |
| RADx-UP Project Data | RADx-UP projects will deposit extant or new data into the *RADx-UP Data Portal.* These data are ingested and harmonized with the RADx-UP Common Data Element specifications and contribute to the RADx-UP de-identified data set. |
| RADx-UP Data Platform | A data platform developed and maintained by the *RADX-UP CDCC* and consisting of two data repositories created from RADx-UP project data. The first repository includes de-identified essential (tier 1) and recommended (tier 2) common data elements as well as additional data provided from RADx-UP |

| | |
|---|---|
| | project awards. The second repository is highly secure and contains the patient re-contact registry and includes participant recontact information for use by RADx-UP investigators who have approved studies requesting re-contacting of RADx-UP participants for future research studies. |

# IV. Policy

## A. Defining Data Quality Categories

Data quality issues may occur in any of the following instances:

1. **General:**
   - Data uploads are not conducted at the intervals agreed upon by the individual project and RADx-UP CDCC.
   - Data upload do not align with the project's DTA/DUA and the informed consent form.

2. **Common Data Elements (CDE)**

   This includes all data submitted in RADx-UP CDE format, include data originally collected in CDE data format as well as data mapped from either previous survey data, electronic health records, or other sources, to the RADx-UP CDE format. Data may fail to meet data quality standards in any of the following ways:

   - Conformance: Uploaded data does not meet RADx-UP data conformance guidelines around variable naming conventions, response options and coding, branching logic, formatting, etc. (See Data Submission Guidance and EHR mapping guidelines on *my*RADx-UP*home*).
   - Actual vs. Expected: The set of CDE variables uploaded does not match the set agreed upon for submission by the project in their NIH CDE Tracking Form; the NIH CDE Tracking Form must align with the outcome letter of the NIH review process.
   - The set of CDEs received fails logic checks.
   - The set of CDEs received fails plausibility checks.

## B. Data Quality Resolution – Triggering the Standard Resolution Process

1. **Data Upload Frequency:** Projects are expected to upload datasets to the RADx-UP Data Portal on a weekly basis. Projects that are not able to complete weekly uploads should discuss this with their EIT and CDCC Informaticist, and set an adjusted schedule of bi-weekly or monthly uploads. However, projects are also not expected to upload datasets at expected intervals if no new data has been obtained by the project since the last data upload.

If a project misses the following number of uploads, the resolution process will be triggered:

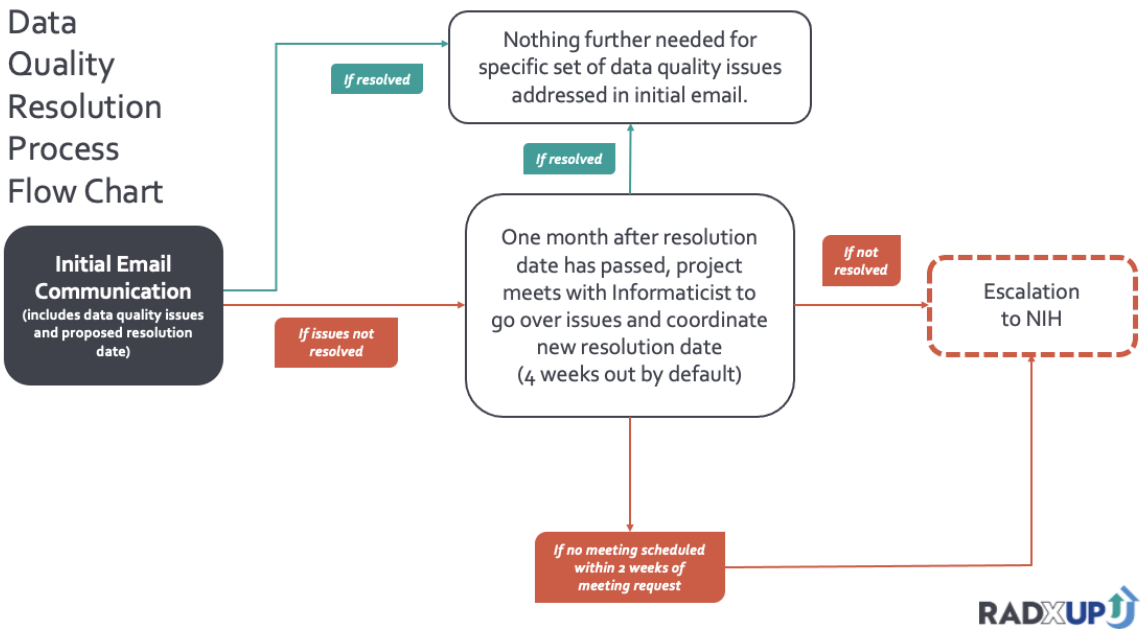
- Scheduled weekly uploads: 4 consecutive uploads are missed (triggered on day 31)
- Scheduled bi-weekly uploads: 2 consecutive uploads are missed (triggered on day 31)
- Scheduled monthly uploads: 1 upload is missed (triggered on day 31)

The resolution process parameters are based on the need for the RADX-UP CDCC to ensure no issues are occurring with dataset uploads and resolution begins for all projects on day 31.

2. **CDE Quality and Conformance:** When identified project uploads data that includes issues with conformance, logic, plausibility, or actual vs. expected data set contents, The *my*RADx-UP*home* portal will generate and display a data quality report to the uploader. This report can also be accessed and referenced at a later date. All issues identified in this report should be addressed within a month. If issues are not addressed, the standard resolution policy will be triggered on day 31.

## C. Data Quality Resolution – Standard Resolution Policy

1. On day 31, the CDCC Informaticist will reach out to the project and copy the EIT. The communication will include details about any relevant issues related to data upload frequency or data quality and conformance (conformance, logic, plausibility, actuals vs. expected). The email will include a request for a meeting between the project team and the informaticist.
    a. During the meeting, the informaticist will review any data issues with the project and help them work through solutions. A timeline will be set for resolving all data quality issues.
    b. If the project resolves all data issues identified in the meeting with the informaticist within the agreed time frame, the issue is resolved. Projects may ask for an extension if necessary.
2. If the project does not resolve all issues within the specified timeline and does not communicate about progress or request an extension of time, the informaticist will send the project an email to let them know that the issue is now being escalated to the NIH. The Informaticist with work with the EIT to set up a meeting with the Informaticist, EIT, project team, Project PO, and RADx-UP PO to discuss the issue.
3. If a project does not set a meeting with the informaticist within two weeks of the request or communicate scheduling delays (such as required staff out of office), project data quality issues will be escalated to the NIH. Meeting does not have to occur within two weeks of the initial email, but should be scheduled within two weeks.




Resolution steps above only apply to outstanding data upload errors that are included in informaticist communications. New sets of errors initiate new resolution timelines.

## D. Data Quality Resolution – Data Upload Not Compliant with DTA/DUA or Informed Consent Form

Projects are expected to upload datasets that are compliant with the terms and conditions in their fully executed Data Transfer Agreement/Data Use Agreement, their IRB-approved Informed Consent Form, and each participant's completed informed consent form. When a project provides data that does not comply with the above listed documentation, this constitutes the CDCC holding data that the project or participant did not approve of transferring. Therefore, the resolution of this issue is on a truncated timeline and requires that projects do not upload additional data until the issues are resolved.

If a project uploads a dataset that is not in compliance with all of these documents, resolution will follow the steps indicated below:

1. The CDCC Informaticist will reach out to the project team and copy the EIT as soon as the issue is identified.
    a. The communication will include details of the issue, instructions, and a 2-week timeline for resolution. Projects will be expected to fully comply with the instructions and timeline. If the project needs additional time to correct the issue, they should request additional time from the informaticist.
    b. No additional data should be uploaded until the issue is resolved as this issue

constitutes the CDCC holding data that the project or participant did not approve of transferring.

2. If resolution is not completed according to the instructions and timeline provided, the issue will be escalated to the project's NIH Program Officer/Project Scientist.

## V. Review and Revision

The RADx-UP Data Core will review and approve updates to this policy as needed.

## VI. Supporting documents

A. RADx UP Data Submission Guidance
B. RADx-UP EHR Data Guidance

## VII. References

## VIII. History of Change

| Section Affected | Version Date | Changes Made |
|---|---|---|
| | | |
| | | |

## Approval Page

| Name/Title | Signature/Date |
|---|---|
| Approved by:<br><br>Warren Kibbe, Ph.D, FACMI<br>MPI, RADX-UP CDCC | Warren A Kibbe, Ph.D.<br>Digitally signed by Warren A Kibbe, Ph.D.<br>Date: 2021.10.25 19:40:42 -04'00' |