AIRA
AMERICAN IMMUNIZATION REGISTRY ASSOCIATION

Data at Rest

Frequently Asked Questions

November 2021

Table of Contents

Basics About Data at Rest....2
How To Get Started....2
DAR Process....3
Benefits of Being Measured....5
Operational Questions....6
Funding and Costs....10
Technical Assistance....11

Basics About Data at Rest

1. What is the Measurement and Improvement (M&I) content area Data at Rest?

The American Immunization Registry Association's (AIRA's) first Measurement and Improvement data quality content area is Data at Rest (DAR). DAR is the measurement of data residing in the IIS production database regardless of how it arrived there. This content area puts into practice data quality indicators found in the May 2018 IIS Data Quality Practices document.

AIRA will be measuring:

- 26 Completeness Indicators
- 23 Validity Indicators
- 8 Timeliness Indicators

2. How is DAR different from other M&I content areas?

Other content areas (i.e. Transport, Submission, Query, Clinical Decision Support) only need credentials provided from IIS staff in order to be measured. DAR will require more involvement from IIS staff to be measured, uses de-identified data from within a production environment instead of test data from within a pre-production environment, and may require a partner agreement between AIRA and the IIS Program.

3. What additional involvement will IIS staff need to provide to be measured?

The IIS team will need to perform three major activities:

- Review the extract specification and determine which data elements can be extracted
- Extract data into the pre-defined file format. This will most likely be performed through SQL scripts. It is possible that your IIS (or some other tool) will be able to extract the data without SQL scripts, but for planning purposes, it would be best to assume SQL scripts will be needed.
- Run the AIRA-provided transformation tool and upload data in AART through the jurisdictional dashboard.

How To Get Started

4. Do I have to have a partner agreement?

No, each jurisdiction will determine the need for a partner agreement with AIRA in order for AIRA to receive the de-identified data. AIRA does not require this agreement; however, some jurisdictions have expressed a need for this type of agreement and AIRA hopes to decrease barriers by providing this document for those jurisdictions that need it.

5. If needed, where do I find the partner agreement?
   AIRA has drafted a partner agreement and is available by reaching out to the AIRA DAR team.

6. What staff would AIRA recommend be involved in this process?
   AIRA would recommend that those currently involved in other Measurement and Improvement activities also be involved in the DAR content area. In addition, others may need to be involved such as central IT, IIS vendors/implementors, or other policy staff.

DAR Process

7. What are the major steps in the process to be measured in DAR?
   The Data at Rest project can be thought of in the context of 6 specific steps.

   - Project Initiation
   - Data Extraction
   - Data Transformation
   - Data Loading
   - Data Analysis
   - Data Quality Improvement (Not required as part of being measured, but will help with improvement)

8. What is the purpose of the Data Extraction step and what is entailed?
   The purpose of this step is to extract all the necessary patient and immunization data from the IIS into a pre-defined flat file format. The following materials are supplied by AIRA to help with the extract.

   - Flat File Format Specification Draft – This format specification defines a patient and immunization extract for all CDC Endorsed data elements
   - Data Extract Specification Draft – This specification defines the birth cohort and other information related to DAR
   - DAR Extract Files:
     i. DAR Extract Job Aid – This job aid can be used to help an IIS determine which data elements can/can't be extracted
     ii. Example Extract Queries – These queries can be used as examples to help your technical staff get started with extracting data
     iii. Example Extract Files – These files can be used as examples to help your technical staff get started with extracting data

9. What is the purpose of the Data Transformation step and what is entailed?
   The purpose of this step is to transform the patient and immunization data into a summary of data quality indicators. The files to support data transformation can be found here under the Data Transform section. These tools were developed by The National Institute of Standards and Technology (NIST).

10. What is the purpose of the Data Load step and what is entailed?
    The purpose of this step is to submit the encrypted aggregate detections file created during the previous step to AIRA for analysis. This file can be uploaded to AART through the jurisdictional dashboard.
11. What is the purpose of the Data Analysis step and what is entailed?
    The purpose of this step is for AIRA to use the NIST provided tool to analyze the aggregate detections file. This step is complete when the data is ready to be displayed in AART.
12. What is the purpose of the Data Quality Improvement step and what is entailed?
    Similar to previous content areas (i.e. Transport, Submission, Query, etc.), this step includes logging into AART, reviewing test results, and determining the best path forward to improve data quality. This step could also include using the NIST tool to dive deeper into your data to find specific areas of concern or submitting provider data where targeted outreach could potentially increase quality of the data that is being sent to the IIS.
13. What materials will AIRA supply to help my IIS in the process?
    AIRA has developed resource materials available on the Data At Rest page to support participation for this content area.
    - DAR Checklist – This checklist is provided to DAR participants to as a guide to help IIS complete the necessary steps
    - DAR Brief Summary Draft – This document is ideal for leadership and those being introduced to the project
    - Flat File Format Specification Draft – This format specification defines a patient and immunization extract for all CDC Endorsed data elements
    - Data Extract Specification Draft – This specification defines the birth cohort and other information related to DAR
    - DAR Extract Job Aid – This job aid can be used to help an IIS determine which data elements can/can't be extracted
    - Example Extract Queries – These queries can be used as examples to help your technical staff get started with extracting data
14. What are the major activities involved through this process?
    - Once a general understanding of the extract concepts and data elements is achieved, the Extract Job Aid is the best document to review next. The job aid will allow for all IIS team members to work from the same artifact to capture the IIS specific nuances. These can then later be used by technical staff to develop the extract (usually using an SQL script).
    - The second major activity in this step is to build and perform the extract into the defined flat file format per the CDC Endorsed Data Element Extract

Specification. The provided Example SQL Functions should be leveraged at this point to help build the extract.

The Data Extract step is finished once both the patient and immunization file have been produced. The next step (Data Transform) will verify the extracted files are formatted properly. It is likely you may need to return to this step and tweak the format and/or data selection based on feedback in the Data Transform step.

Benefits of Being Measured

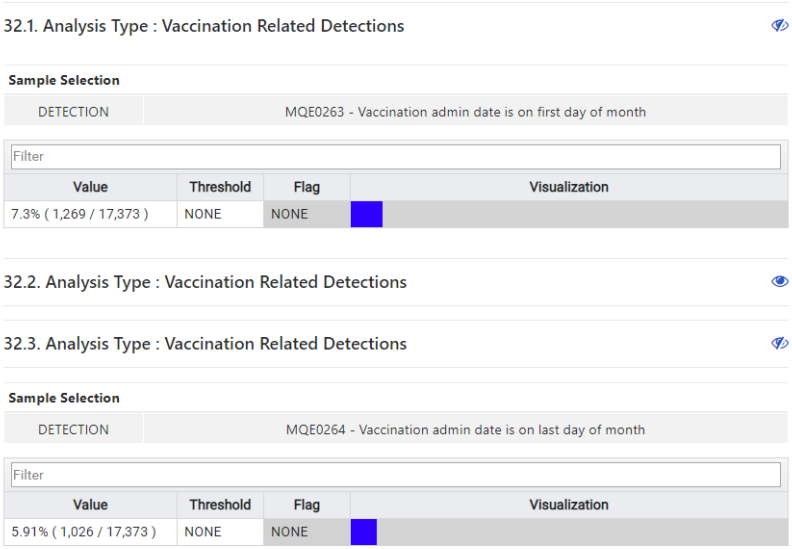
15. What are the anticipated benefits of being measured as part of the M&I content area DAR?

Measurement and Improvement provides IIS with an objective, independent testing process; the available data will enable individual IIS to share their progress toward meeting standards with the broader community. This in turn will build confidence in IIS across other communities and will help prioritize essential improvements and community-wide adoption of standards across both IIS and electronic health record (EHR) systems. This content area will specifically help IIS to understand their data quality within the data that is "resting" in their data base. This content area puts into practice data quality indicators found in the May 2018 IIS Data Quality Practices document.

AIRA will be measuring:

- 24 Completeness Indicators
- 18 Validity Indicators
- 2 Timeliness Indicators

Results from this content area can be used to highlight a potential data entry problem. For example, the image below is highlighting a potential data entry problem for this IIS where data is heavily skewed to the 1st and last day of the month. This may or may not indicate an actual data quality problem, but it is statistically unlikely that 13.2% of all records (7.3% on 1st and 5.9% on last) were administered on two days of the month.



16. What about IIS who are not yet ready to be measured in the DAR content area?

Measurement is encouraged but remains voluntary. The AART application will continue to be used as the platform for declaring a desire to be measured in the stages and content areas of M&I, as well as for declaring sharing settings to allow for transparency. To lower the burden, each IIS will be rolled forward to be included in measurement and share in new content areas and stages; however, AART users can change measurement and sharing settings at any time, and "no" responses will be rolled forward as well.

17. Is technical assistance available if my IIS is not sure how to get started or needs help along the way?

Technical assistance is available to IIS Programs and Vendors as needed to support all measurement and improvement efforts. IIS are encouraged to be measured even if they won't meet specific measures.

Operational Questions

18. Who will oversee the M&I DAR process?

All measures and tests throughout the Measurement and Improvement effort have been and will continue to be developed by MACAW, vetted for comment with the full IIS community via a Town Hall meeting webinar format, and reviewed and approved by the AIRA Board of Directors prior to implementation.

19. Who will perform the testing?

AIRA staff will conduct all testing. If, at some future point, testing needs exceed AIRA's capacity, alternatives will be explored with MACAW and the AIRA Board, in conjunction with community input.

20. How will the testing be conducted?

AIRA technical staff are responsible for implementing and conducting all testing efforts within the M&I Initiative. For DAR, the testing methodology involves IIS sending aggregate data quality indicators from IIS production systems through secure methods. The aggregate data quality indicators are void of any identifiable information and is a summary of IIS data quality rather than individual record level information. NIST is partnering with AIRA in the testing process to develop conformance test tools in support of this content area.

21. What data indicators are measured in DAR?

The following Data Quality Indicators are being checked as part of the Data at Rest project. These are broken in to four quality dimensions as defined by the IIS Data Quality Practices to Monitor and Evaluation Data at Rest document found here.

- **Completeness:** The degree to which full information about a data set, record, or individual data element is captured in the IIS, i.e., the proportion of stored data measured against the potential of "100%."
- **Validity:** The degree to which the data conforms to the syntax (format, type, range) of its definitions, i.e., to the rules of what is accepted or expected by the IIS.
- **Timeliness:** The amount of time between the occurrence of the real-world event and its documentation in the IIS, i.e., the time lag between the date of vaccination or birth and the date the record is received by the IIS.

22. What data elements are measured?

The following fields are measured for completeness.

- Patient First Name
- Patient Middle Name
- Patient Last Name
- Patient Date of Birth
- Patient Gender
- Patient Address: Street
- Patient Address: City
- Patient Address: State
- Patient Address: Zip
- Patient Address
- Next of Kin First Name
- Patient Guardian First Name
- Next of Kin Last Name
- Patient Guardian Last Name
- Mother's Maiden Name
- Race
- Ethnicity

- Phone Number
- Email
- Vaccine Administration Code
- Vaccine Administration Date
- Vaccine Information Source
- Vaccine Lot Number
- Vaccine Lot Expiration Date
- Vaccine Eligibility
- Vaccine Funding Source

As noted in the Data at Rest document, Validity and Accuracy dimensions are closely related and often measured using a shared set of indicators. As such, they were combined in the Data at Rest Document and combined here.

- 1st, 15th, and end of month Date of Birth
- Birth Date is over 120 years ago
- Administration Date is after Lot Expiration Date
- Administration Date is before Birth Date
- Administration Date is before/after the minimum/maximum age for routine vaccine age ranges
- More vaccines than expected for patient age
- 1st, 15th, and End of Month Vaccination Dates
- Administered vaccines with minimal details (e.g., might actually be historical)
- Lot numbers that violate lot number patterns.
- Administered vaccines with unspecified vaccine codes
- Unrecognized CVX Codes
- Eligibility and Funding Source Disagree
- Unrecognized Route
- Unrecognized Site

The following measure the amount of time it took for the event (birth or vaccination) to enter the registry by comparing the event date with the system creation date.

- Birth Record Creation
- Vaccination Record Creation

23. What types of reports will be available?

- Similar to all other content areas (e.g., transport, Submission), each IIS will receive an individualized report with in AART.
- Each IIS will have control over who – if anyone – is allowed to see their report in either named or unnamed fashion
- Each IIS will also have access to additional reports based on their data on the NIST tool, qDAR.

- Once Data at Rest moves to the Assessment stage, AIRA will develop aggregate reports based on all IIS who submit data
- This will be a once-a-year activity as opposed to quarterly like other content areas.

24. What results will be shared?

AART Admin users can set sharing settings in AART. These sharing settings continue to function as they do in other content areas (i.e. Transport, Submission, Query, etc.)

25. How often will data need to be pulled and sent to AIRA?

We anticipate receiving, measuring, and providing results for the DAR content area once per year.

26. How much time is needed from IIS staff for developing the extract?

That amount of time will vary by IIS. However, after training and the initial data pull, we do not anticipate this taking more than active 5 hours of staff time for each extract. Reviewing and understanding the results after measurement will take more IIS staff time and will be at the discretion of each IIS. IIS should plan on the first extract taking more time.

27. How much time is needed from IIS staff for the whole process?

During the pilot implementation of DAR, AIRA found that IIS ranged from 40 to 80 hours of time to perform the entire process from start to finish. The amount of time will depend on each IIS.

28. How much data needs to be included each time I send data to AIRA?

The DAR extract is comprised of a cohort of patients born in the past two calendar years prior to submission and their vaccination information.

29. Are there computer minimum specifications to run the tool?

Yes, the only item the computer will need is a recent version of Java. No other specifications are required to use the tool provided by NIST.

30. Do you have a statement of assurance or validation of some kind for the software that I can provided to my IT department to support installation and use?

Yes, AIRA has worked with NIST and can provide a technical statement describing what the NIST tool does and how it works. In addition, documentation to support the use of NIST tools can be found here in the Data Transformation step.

31. Can an IIS verify what data is in the encrypted detections file that will be sent to AIRA?

Yes, NIST has created a way for the IIS to see what the encrypted detections file entails prior to sending it to AIRA.

32. Will the NIST tool need updates or will IIS need to get newer versions?
    We do not anticipate frequent changes to the NIST data at rest tool. However, updates may occur and AIRA would provide the update tool to IIS prior to the measurement timeframe.

33. Does any protected health information leave the jurisdictions IIS?
    No, as with all content areas within M&I, AIRA is careful to not include any identified data. AIRA worked closely with NIST on a solution that only sends numbers of data quality detections without any identifying information.

34. How long will the data I am sending to AIRA be retained and will we receive confirmation of it being destroyed?
    Currently, all data is hosted on a secured Microsoft Azure cloud. If AIRA's retention policy changes in the future, IIS will be notified of the timeline and process.

35. Who can access the data that is sent to AIRA?
    Only AIRA technical staff have access to data from M&I content areas. We do periodically receive information requests from CDC and partners. We honor all sharing settings when responding to these requests, and mostly provide aggregate data across all IIS.

36. What will the data be used for?
    The data will only be used in the NIST tool to determine detections for data at rest. Results from the analysis of this data will be displayed in AART. Settings within AART will function the same as they do with other content areas being measured as part of M&I (Transport, Submission, Query, etc.)

37. Could our jurisdictional firewalls or other permissions cause challenges?
    Yes, if the user attempting to use the NIST provided tool does not have appropriate permission, this could be a barrier. However, we anticipate that for most jurisdictions the use of this tool will fall within normal day-to-day computer use for IIS staff.

38. If I am cloud hosted, can my IIS vendor send AIRA the data?
    Yes, AIRA will work with all IIS vendors and implementors as needed.

Funding and Costs

39. What will it cost my IIS to be measured?
    All testing is conducted by AIRA at no cost to Immunization Programs and IIS. Recommended system enhancements discovered through the testing process may have system modification costs. IIS will have other resource costs associated with participating in M&I; however, as independent 3rd party testing increases, costs associated with other self-reporting measures (e.g., IIISAR) may decrease.

40. Who is funding AIRA to do this testing work?

The M&I Initiative is part of AIRA's Standards Cooperative Agreement with CDC. The Cooperative Agreement charges AIRA with conducting evaluations of IIS adherence to standards and identifying gaps in adoption and implementation.

As a member organization, AIRA's primary mission is to serve our members. The AIRA Board of Directors guides the design and implementation of initiatives.

Technical Assistance

41. Is there technical assistance available to help my IIS staff, Immunization Program Manager, IIS Vendor, or Central IT partners prioritize and plan enhancements?

Yes, AIRA technical and programmatic staff are available to meet with teams to strategize on how best to prioritize enhancements.

42. Where can I find more information about the Measurement and Improvement process?

Please visit the AIRA website for more information on the Measurement and Improvement Initiative.