



AIRA

AMERICAN IMMUNIZATION
REGISTRY ASSOCIATION

Data at Rest

Extract Specification

Updated January 2025

Table of Contents

Introduction	2
Overview	2
<i>Figure 1: Extract process</i>	2
Output Format for Extracted Data Elements	2
Patient and Vaccine Extract File Format	3
<i>Figure 2: Extraction Decision Tree</i>	4
Required elements.....	5
Patient data elements	6
Extract Cohort	6
Patient Extract File	8
Vaccine Extract.....	12
Change History	17

Introduction

This extract specification describes the data elements necessary to extract from an IIS for Data at Rest. This extract specification is built from the CDC [Flat File Format Specification](#). Reviewing and understanding that specification may be important to successfully formatting the files, but it serves only as a reference. Readers participating in the Data at Rest content area can instead refer to the patient and immunization data elements below. Selected data elements will be coded as excluded.

Note that there are two separate extract files – one for patients and one for immunizations. These files are linked together with the patient’s unique ID.

Overview

Key steps (refer to [Figure 1](#)):

1. Develop/implement process to extract specified data elements
2. Execute extraction process (e.g., DBMS queries)
3. Store the extracted patient data elements and vaccine data elements into a separate ASCII text files (tab separated values) with a unique IIS Patient ID used to associate data between them.

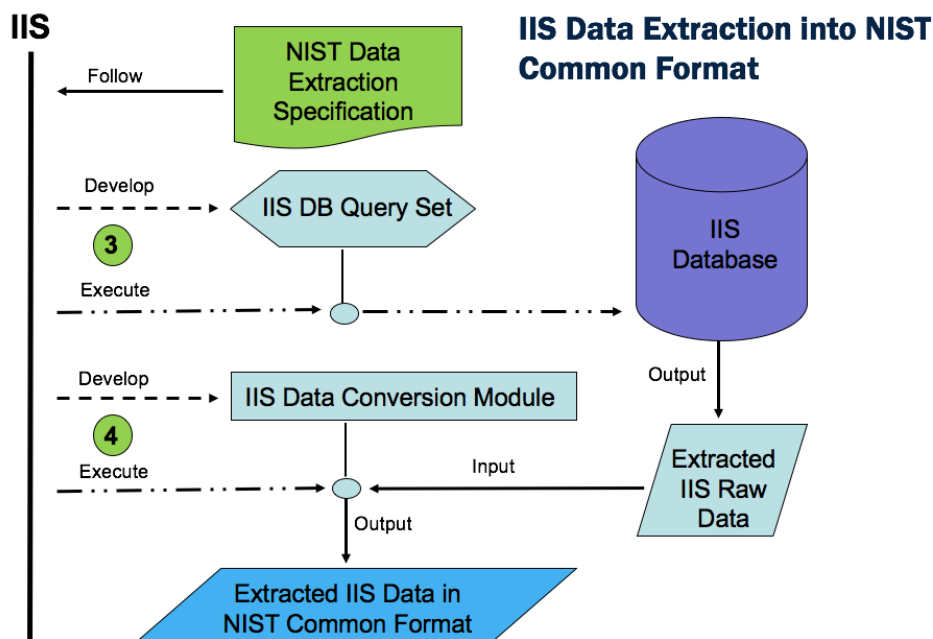


Figure 1: Extract process

Output
Format for
Extracted
Data
Elements

All data shall be output to the relevant Extract File as UTF-8 encoded text strings¹ without

¹ <https://en.wikipedia.org/wiki/UTF-8>

change from the local representation except for:

- Elements that represent Dates (e.g., date of birth) shall be stored using the date string format: *YYYY-MM-DD (4-character year; 2-character month; 2-character day)*.
- Elements represented using a coding system (e.g. Product Type may use CVX codes, local codes, etc.) shall be stored according to the coding system prescribed in HL7 v2.5.1.

Patient and Vaccine Extract File Format

Extracted patient and vaccine-related data elements (see [Table 1](#) & [2](#)) shall be stored in two separate files, a Patient Extract File and a Vaccine Extract File. A Patient ID data element is used to link every patient record to its corresponding vaccine record(s). The structure of both files, referred to below simply as “Extract File,” are identical.

Each Extract File is a text file containing lines of tab² separated values (TSV), the first line may consist of Data Element Name values from the tables below (Patient or Vaccine) (i.e., the column headings or may be the first line of data with no column headings). Each subsequent line represents a logical “record,” consisting of tab separated values containing the extracted values corresponding to each “column.”

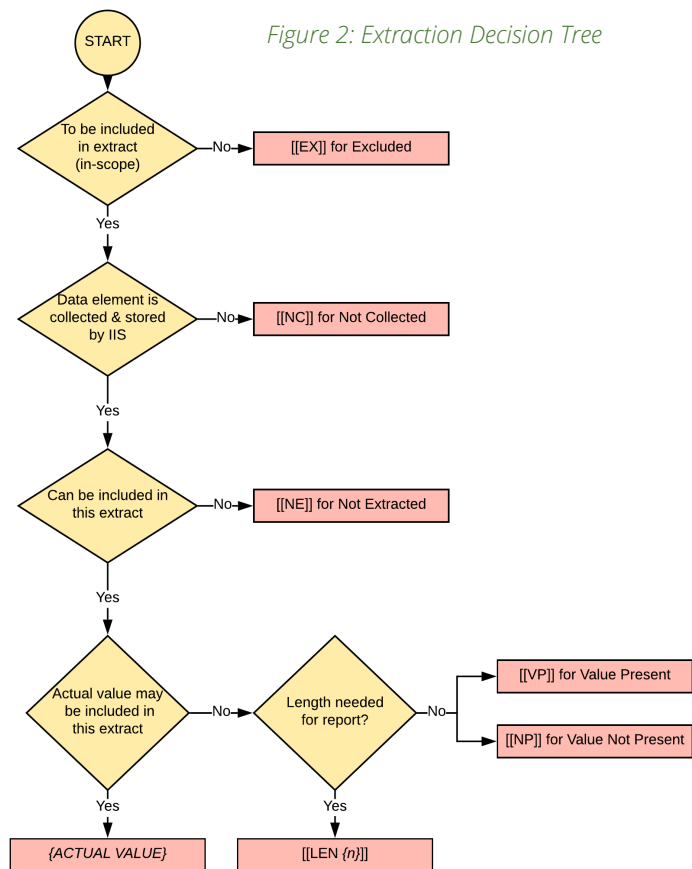
Note that it is possible for some extracted data elements themselves to contain one or more tab (ASCII 09) characters. In such instances, any tab characters *appearing in data elements* must be mapped to space (ASCII 32) characters in the Extract File.

Prior to data extraction, the IIS should thoroughly review this document to define which patient and vaccine-related data elements are “in scope” and should be extracted versus those elements which should not be extracted (i.e., excluded). Additionally, prior to data extraction, the IIS shall determine which of the above “in scope” elements may or may not be extracted for any of the following reasons:

- the data element is never collected/stored by the IIS
- local policies prohibit release of the element, and any metadata associated with it (ex. presence, length, etc.)
- local policies only permit release of metadata only

² “tab” refers to the ASCII HT (Horizontal Tab) character, which has a decimal value of 09, and is the same character typically generated by pressing the ‘Tab’ key on most keyboards.

The determination of which data elements shall or shall not be extracted must be consistent across all lines (i.e., records) of the Extract File. Thus, each data element shall either *always* be extracted (and therefore output to the Extract File), or *never* extracted, in which case a predefined string indicator (see [Figure 2](#)) shall be output to the Extract File in the respective column for that element. Figure 2 provides a flow chart of how to determine what should be output to the Extract File for any given data element.



Step 1: the IIS consults this document to determine which elements are “in scope” for the extraction. Elements not in scope shall be represented by the special string “[[EX]]” for excluded (quotes not included) in the Extract File.

Step 2: the IIS determines which of the “in scope” data elements are extractable (i.e., are they in the IIS database?). Elements that cannot be extracted because they are not collected and/or not stored in the IIS database shall be represented by the special string “[[NC]]” for Not Collected (quotes not included) in the Extract File.

Step 3: IIS determines, based on their local policy, which of the remaining data elements (after steps 1 & 2 above) they are *willing* to extract. For example, the IIS may choose *not* to extract the element out of privacy concerns. In that case, the element shall be represented by the special string “[[NE]]” for Not Extracted (quotes not included) in the Extract File.

Step 3 (Variation A): Some IIS, however, may determine that although they are not willing to extract the actual values of certain data elements, they *are* willing to indicate whether or not the element is or is not present in the IIS database (such

information is essential to the quality analysis for measurement of data “completeness”). In such cases, the special strings “[[VP]]” for Value Present and “[[NP]]” for Value Not Present (quotes not included) may represent the element in the Extract File.

Step 3 (Variation B): Similarly, some IIS may determine that although they are not willing to extract the actual values of a particular data element, they *are* willing to extract its length (e.g., it may be useful to the quality analysis to know the length of some data elements such as Patient Name). In such cases the special string “[[LEN {n}]]” (quotes not included) shall be used to represent the data element in the Extract File, where {n} is an integer representing the string length of element’s value in the IIS database. Note that data elements that contain empty/null (ex. note present) values in the database should result in {n} being set to 0.

Note that the above cases where data elements *are not* extracted is different from the cases where data elements *are* being extracted but contain empty/null (ex. not present) values in the IIS database. The latter shall result in <tab><tab> being output to the Extract File.

Required elements

Only data elements necessary to link records together (e.g., IIS Patient ID) or are used for report creation (e.g., Reporting Group) are required to always be populated. These are noted in bold. All other fields should be populated or left empty as indicated by the rules above and the directions for this particular export.

The linking elements are required to be unique between all associated files and for the purposes of this extract. However, they are not required to represent the actual identifier, as used by the IIS. They may be generated for one-time use during the extraction. In other words, a record linked with IIS Patient ID “100” may or may not represent the same record in an export at a later date with the same IIS Patient ID “100”. The linking elements are only required to be unique and consistent across a single set of files extracted and submitted at the same time. Thus, extractors have the option of further obfuscating the original patient id by not being required to export usable patient identifiers.

For all other fields, extractors must not skip records that are missing data that would normally be required by an IIS or ensure that empty values are populated with contrived information. This specification is designed to allow extractors to represent the data as it is, with the least amount of translation.

Patient data elements

The patient-related data elements to extract to the Patient Extract File are listed in the table below.

Note that all of the patient records in the Patient Extract File shall have at least one record in the Vaccine Extract File (see next section) unless the patient has no immunizations. Thus, if a patient record does not have a corresponding immunization record the analysis will assume this patient has received no immunizations.

Also note that patient information is assumed to be “consolidated” by the IIS prior to extraction, and thus a single set of the elements listed below are expected to be extracted for each Patient. Some IIS may store multiple representations of the same patient record or vaccination event, care should be taken to extract the single best representation (as the IIS can determine it) of the elements below. IIS should not extract all versions and variations of received and stored records.

Extract Cohort

The extract should include all patients (and all their associated immunization records) provided they meet all the following criteria:

1. The patient’s DOB is on or after January 1, 2023 and on or before December 31, 2024.
2. The patient meets one of the following:
 - The patient is known to reside in the jurisdiction.
 - The patient’s address is unknown, and the IIS has never received vaccination information for the patient.

The patient extract should include all patients provided they meet all the following criteria:

1. The patient’s date of birth (DOB) is on or after January 1, 2023, and on or before December 31, 2024.
2. The patient meets one of the following:
 - The patient is known to reside in the jurisdiction.
 - The patient received an immunization from a provider organization within the jurisdiction, but the address is unknown.

The immunization extract should include all immunizations for those in the patient extract.

This patient extract should NOT include:

1. Patients who are not active at the jurisdictional level (e.g., inactive, deceased).
2. Patients who do not consent to be in the registry or whose parents do not consent to share their records.
3. Known test patients.

Patient Extract File

This extract further defines the values and data selection expectations from the base Data at Rest Extract Specification. The resulting file shall meet the technical requirements defined in the base document.

Table 1: Patient Data Elements

	Data Element Name	Format	Definition	Extract Guidance	Value Set
1	IIS Patient ID	String	Unique ID for this patient within this extract file.	This may or may not be the actual IIS Patient ID within the IIS DB. A unique value is required within this extract file.	
2	Patient Name – First	String	The patient’s first name.	Populate with either [[VP]] (Value Present) or [[NP]] (Value Not Present) Note: for this field and others, IIS are to use the above codes as literally written – a two-letter code surrounded by double brackets.	
3	Patient Name – Middle	String	The patient’s middle name.	Populate with either [[VP]] or [[NP]]	
4	Patient Name – Last	String	The patient’s last name.	Populate with either [[VP]] or [[NP]]	
5	Mother’s Maiden Name	String	The last name under which the mother was born (i.e., before marriage).	Populate with either [[VP]] or [[NP]] This should be the mother’s maiden name and not inferred or implied from a responsible person value unless it is explicitly known to be the mother’s maiden name.	

6	Mother's Name – First	String	The first name of the patient's mother.	Populate with either [[VP]] or [[NP]] This can either be from an explicit Mother's name field or a responsible person where the relationship to the patient is "Mother".
7	Mother's Name – Middle	String		Default this field to [[EX]] (Excluded) for all records.
8	Mother's Name – Last	String	The last name of the patient's mother.	Populate with either [[VP]] or [[NP]] This can either be from an explicit Mother's name field or a responsible person where the relationship to the patient is "Mother".
9	Patient Date of Birth	YYYY- MM-DD	The patient's date of birth.	Populate with actual value from the IIS production database (DB).
10	Patient Gender	String	The patient's gender.	Populate with either [[VP]] or [[NP]]
11	Patient Address – Street	String	The street component of the patient's address.	Populate with either [[VP]] or [[NP]] This should be the address the IIS considers to be the most accurate address for the patient.
12	Patient Address – City	String	The city component of the patient's address.	Populate with either [[VP]] or [[NP]] . This should be the address the IIS considers to be the most accurate address for the patient.
13	Patient Address – State	String	The state component of the patient's address.	Populate with either [[VP]] or [[NP]] This should be the address the IIS considers to be the most accurate address for the patient.
14	Patient Address – Country	String		Default this field to [[EX]] for all records.
15	Patient Address – ZIP/postal	String	The ZIP code of the patient's address.	Populate with either [[VP]] or [[NP]] This should be the address the IIS considers to be the most accurate address for the patient.

16	Patient Race	String	The patient's race.	Populate with either [[VP]] or [[NP]]
17	Patient Ethnicity	String	The ancestry of the patient.	Populate with either [[VP]] or [[NP]]
18	Patient Telephone Number	String	The patient's telephone number.	Populate with either [[VP]] or [[NP]] This can be a telephone for the patient or one of their responsible persons.
19	Patient Email Address	String	The patient's email address.	Populate with either [[VP]] or [[NP]] This can be an email address for the patient or one of their responsible persons.
20	Patient Primary Language	String		Default this field to [[EX]] for all records.
21	Patient Alias Name: First	String		Default this field to [[EX]] for all records.
22	Patient Alias Name: Middle	String		Default this field to [[EX]] for all records.
23	Patient Alias Name: Last	String		Default this field to [[EX]] for all records.
24	Responsible Person Name: First	String	The first name of the person responsible for the patient.	If at least one responsible person with a first name exists, populate with [[VP]] , otherwise populate with [[NP]]
25	Responsible Person Name: Middle	String		Default this field to [[EX]] for all records.

26	Responsible Person Name: Last	String	The last name of the person responsible for the patient.	If at least one responsible person with a last name exists, populate with [[VP]] , otherwise populate with [[NP]] .
27	Responsible Person Name: Relationship to Patient	String		Default this field to [[EX]] for all records.
28	Patient Multiple Birth Indicator	String		Default this field to [[EX]] for all records.
29	Patient Birth Order	String		Default this field to [[EX]] for all records.
30	Patient Birth State	String		Default this field to [[EX]] for all records.
31	Patient Status Indicator—Provider Level	String		Default this field to [[EX]] for all records.
32	Patient Status—Jurisdiction Level	String		Default this field to [[EX]] for all records.
33	Record Creation Date	YYYY-MM-DD	The date this patient was created in the IIS.	Populate with actual value (no time component) from DB.
34	Record Update Date	YYYY-MM-DD		Default this field to [[EX]] for all records.

Vaccine Extract

The vaccine-related data elements to extract to the Vaccine Extract File are listed in the table below.

Note that all immunization records in the Vaccine Extract File shall have a corresponding patient record in the Patient Extract File ([see section above](#)). If an immunization record does not have a corresponding patient record in the immunizations will be noted and discarded from analysis.

Table 2: Vaccination Data Elements

	Data Element Name	Format	Definition	Extract Guidance	Value Set
1	IIS Patient ID	String	This is the IIS Patient ID from the Patient Extract File.		
2	IIS Vaccination Event ID	String	Unique ID for this vaccination event within this extract file.	This may or may not be the actual IIS Vaccination Event ID within the IIS Database. A unique value is required within this extract file.	
3	Reporting Group	String	This field will be used to group like submitted immunization records for analysis and reporting purposes.	For DAR, this value should represent the entity that is viewed as the owner of this vaccination event (likely viewed as similar to the Responsible Organization in HL7). This value will be some form of an identifier that makes logical sense to the IIS performing the extract.	

4	Sending Organization	String		Default this field to [[EX]] for all records.	
5	Responsible Organization	String		Default this field to [[EX]] for all records.	
6	Administered at Location	String	The facility name/identifier of the facility that administered the immunization.	Populate with either [[VP]] or [[NP]] .	
7	Vaccine Administering Provider	String	The name of the provider (person) administering the vaccination. Note that this "provider name" could include an actual name, or an ID.	Populate with either [[VP]] or [[NP]] .	
8	Vaccine Type (CVX)	String	The vaccine type of the vaccination event.	Populate with the CVX corresponding to the vaccination event. This can either be a direct extract or a translation if CVX codes are not stored directly with the vaccination event record. It is expected IIS will use leading zeros on two-digit CVX codes (e.g. "03" for MMR).	CVX
9	Vaccine Type (NDC)	String	The vaccine type of the vaccination event.	Populate with the NDC corresponding to the vaccination event. This can either be a direct extract or a translation if NDC codes are not stored directly with the vaccination event record. This	NDC

				must be the actual NDC used during recording the vaccination event or it should be left blank. Do not populate with a possible NDC if one was not used during recording the vaccination event.	
10	Vaccine Administration Date	YYYY-MM-DD	The date the vaccination event occurred.	Populate with the actual value from the DB.	
11	Vaccine Manufacturer	String	The manufacturer of the vaccine that was administered.	Populate with the MVX. This could have been received directly as an MVX code or translated from an NDC. If neither were submitted, then this field should remain blank.	MVX
12	Vaccine Lot Number	String	The lot number of the vaccine administered.	Populate with the actual value from the DB.	
13	Vaccination Event Record Type	String	Indicates whether the vaccination event is based on a historical record or was given by the administered at location.	Populate with table NIP001 value (e.g., "00" for new immunization record, "01" for historical information - source unspecified, etc.).	NIP001
14	Vaccine Route of Administration	String	The route of administration.	Populate with the either HL7-defined table 0162 or NCIT value when route is present.	HL7-defined table 0162 or NCIT
15	Vaccine Site of Administration	String	The anatomical site where the vaccine was given.	Populate with HL7-defined table 0163 value when site is present.	HL7-defined table 0163

16	Vaccine Expiration Date	YYYY-MM-DD	The expiration date of the vaccine administered.	Populate with actual value from the DB.	
17	Vaccine Dose Volume	String	The amount of vaccine administered (e.g., 0.5).	Populate with either [[VP]] or [[NP]] . Note all volumes will be assumed to be mL. If actual mL are not captured in the IIS, then this field should be set to [[NC]] (Not Collected) for all records.	
18	Ordering Provider	String	The identifier of the provider (person) ordering the immunization. Note: this "provider name" could include an actual name, or an ID. May also be referenced as the prescriber ordering provider Identifier.	Default this field to [[EX]] for all records.	
19	Completion Status	String	Indicates if the vaccination event was successfully administered.	Default this field to [[EX]] for all records.	HL7-defined table 0322
20	Dose Level Eligibility	String	The program that should pay for a given immunization, based on the characteristics of the patient and the type of vaccine administered.	Populate with User-defined table 0064 table value.	User defined table 0064

21	Vaccine Funding Source (Dose Level Public/Private Indicator)	String	The funding source of the vaccine administered, (i.e., was the vaccine administered purchased as publicly funded, privately funded, or with other jurisdiction specific funding.)	Populate with PHVS_ImmunizationFundingSource_IIS value	PHVS_ImmunizationFundingSource_IIS (2.16.840.1.114222.4.11.3287)
22	VIS – Identifier	String	The Vaccine Information Statement (VIS) GDTI document code	Default this field to [[EX]] for all records.	
23	VIS – Publication Date	YYYY-MM-DD	The publication date of the VIS.	Default this field to [[EX]] for all records.	
24	VIS – Given Date	YYYY-MM-DD	The date the VIS information was provided to the patient.	Default this field to [[EX]] for all records.	
25	Record Creation Date	YYYY-MM-DD	The date this vaccination event was created in the IIS.	Populate with actual value (no time component) from DB.	
26	Record Update Date	YYYY-MM-DD	The date this vaccination event was last updated in the IIS.	Default this field to [[EX]] for all records.	

Change History

Version	Date	Author	Change
1.2	1/31/2025	D. Ferlazzo	<ol style="list-style-type: none"> 1. Updated birth cohort for annual testing. 2. Added content from CDC Endorsed Data Element Extract specification. 3. Clarified the patient inclusion criteria. 4. Clarified that leading zeros should be included in applicable two-digit CVX codes (e.g., "03" for MMR). 5. Removed definitions for data elements that should be hard coded as [[EX]]. 6. Requested City and ZIP Code be populated with either [[VP]] or [[NP]] rather than the actual value from the DB. 7. Minor editorial changes to all sections. 8. Minor table reformatting to applicable figures.
1.1	11/30/2023	E. Larson	<ol style="list-style-type: none"> 1. Updated birth cohort for annual testing. 2. Added a Table of Contents.
1.0	10/27/2022	E. Larson	<ol style="list-style-type: none"> 1. Updated birth cohort for annual testing. 2. Field 18 (Patient Telephone Number) Added additional extract guidance. 3. Field 19 (Patient Email Address) Added additional extract guidance. 4. Field 24 (Responsible Person Name: First) Changed extract from defaulting all records with a value of [[EX]] to using either the value [[VP]] or [[NP]] with extract guidance. 5. Field 26 (Responsible Person Name: Last) Changed extract from defaulting a value of [[EX]] to using either the value [[VP]] or [[NP]] with extract guidance.
0.4	11/01/2021	A. Diallo	Birth cohort updated for annual testing.
0.3	3/9/2020	E. Larson	Field 18 (Ordering Provider) in the immunization file was missing Extract Guidance. Added that this field should be populated with [[EX]] for all records.

0.2	02/24/2020	E. Larson	Updated keyword syntax to new shorter format (e.g., changed [[EXTRACTED]] to [[EX]])
------------	------------	-----------	--