



IMPORTING LEGACY DATA TO IMPROVE IIS SATURATION

**RECOMMENDATIONS
AND IMPLEMENTATION
CONSIDERATIONS 7.30.2019**



AIRA
AMERICAN IMMUNIZATION
REGISTRY ASSOCIATION

EXECUTIVE SUMMARY

All immunization information systems (IIS) are affected to some degree by incomplete patient populations and vaccination histories.

A recent survey of the IIS community performed by the American Immunization Registry Association (AIRA) revealed that 57% of IIS programs believe their IIS to be moderately or significantly impacted by incomplete data saturation as compared to 24% that reported a minimal impact.¹ These gaps in data may also become apparent when comparing IIS data against data used to inform the National Immunization Survey (NIS), examining IIS denominators versus Census-based estimates, and investigating discrepancies between records in a provider's electronic health record (EHR) system and records in the IIS. Incomplete patient populations and vaccination histories in the IIS compromise the IIS's ability to provide clinical decision support, inform population-based coverage assessments, support disease outbreak response efforts, and advise policy and resource allocation decisions.

IIS programs spend a majority of their time and effort working with providers to facilitate and improve the timely reporting of newly administered vaccinations; however, the pursuit of older, legacy data is often discounted as programs struggle to balance competing priorities. Legacy data loads can have a direct impact on improving IIS data saturation by increasing the amount of complete patient and vaccination data available in the IIS. More complete vaccination records improve vaccine forecasting, prevent duplicate doses, and decrease missed opportunities to vaccinate. For the purposes of this document, "legacy data" is a term used to describe all patients and all historical and historically administered vaccinations that may be known to a provider or data partner. The term "data saturation" is used to describe the degree to which an IIS has collected all patient and vaccination records within the jurisdiction with the ultimate goal of achieving complete data saturation.

AIRA leveraged a community-informed approach to examine current methodologies for the collection of legacy data, as well as data gap monitoring and resolution processes. These activities were used to inform the development of best practice guidance for standardizing and improving data backloads to increase IIS data saturation. This document is intended for practical use by IIS program staff, EHR vendors, and immunization providers. Information for this project was gathered through an AIRA membership survey, interviews with select IIS programs and EHR vendors, and review of existing documents and resources.

¹ The remaining 19% did not know the extent to which their IIS was impacted by incomplete data and entered a response of "unknown."

IIS programs cited competing priorities, poor provider participation, and concerns about data quality as the primary barriers for pursuing legacy data. Although challenges exist, the value of complete records in the IIS should outweigh any concerns. **IIS programs should ultimately strive to obtain complete immunization records for the entire population served by the IIS.** The IIS community should also continue to develop consistent messages and activities for promoting the value of an IIS and the importance of reporting both newly administered and legacy vaccination data.

This document examines a variety of factors associated with the pursuit of legacy data and offers a series of implementation considerations and recommendations based on input from subject matter experts (SMEs). Topics covered in this document include:

- Optimal timing for pursuing legacy data
- Preferred formats for identifying and transmitting legacy data
- Common data quality and IIS performance concerns
- Identification of alternative data partners
- Strategies for ongoing interface and data monitoring to identify gaps and changes in provider reporting practices
- Suggestions for resolving data gaps and working with providers to correct issues and resubmit missing data to the IIS

An IIS offers its greatest value when fully populated with patient and vaccination data that can be leveraged by all stakeholders that rely on this information. Improving data saturation and partnering with providers to optimize record sharing benefits the entire community and promotes the core public health mission of preventing and containing vaccine-preventable disease.



TABLE OF CONTENTS

EXECUTIVE SUMMARY	i		
1 INTRODUCTION	1		
Background	1		
Audience, methodology, and document organization	3		
2 GUIDANCE, CONSIDERATIONS, AND RECOMMENDATIONS	4		
Legacy import methodologies	5		
HL7 (single-patient protocol)	5		
HL7 batch (multi-patient protocol)	8		
Delimited flat file	11		
Direct data entry	12		
Timing of legacy imports	14		
Validation of legacy records	17		
Data quality concerns	20		
Legacy data creates duplicate patient and vaccination records	20		
Duplicate patients	21		
Duplicate vaccinations	24		
Historically administered doses do not include all currently required fields	25		
Legacy data includes inactive CVX and unspecified vaccine types	27		
Legacy data includes generic vaccinations dates and questionable data sources	29		
		Performance	30
		Program priorities and eliminating barriers	31
		Alternative data sources	34
		Gap monitoring	35
		Gap resolution	36
		CONCLUSION	38
		APPENDIX A: ABBREVIATIONS/ACRONYMS	40
		APPENDIX B: SUMMARY OF IMPLEMENTATION CONSIDERATIONS AND RECOMMENDATIONS	41
		APPENDIX C: REFERENCES	43
		APPENDIX D: ACKNOWLEDGMENTS	44

SECTION 1 INTRODUCTION

1.1 BACKGROUND

Immunization information systems (IIS) facilitate ongoing efforts to gather complete and accurate patient and vaccination information for all residents within their respective jurisdictions. As a result, IIS programs spend considerable time and effort working with providers to facilitate and improve the timely reporting of newly administered vaccinations. In many jurisdictions, the pursuit of older, legacy data is often discounted as IIS programs struggle to balance other competing priorities.

For the purposes of this document, “legacy data” is a term used to describe all patients and all historical and historically administered vaccinations that may be known to a provider or data partner. The term “data saturation” is used to describe the degree to which an IIS has collected all patient and vaccination records within the jurisdiction with the ultimate goal of achieving complete data saturation.

All IIS are affected to some degree by incomplete patient and vaccination data saturation. A recent AIRA survey revealed that 57% of IIS programs believe their IIS to be moderately or significantly impacted by incomplete data saturation, as compared to 24% that reported a minimal impact.² These gaps in data may also become apparent when comparing IIS data against data used to inform the National Immunization Survey (NIS), examining IIS denominators versus Census-based estimates, and investigating discrepancies between records in a provider’s electronic health record (EHR) system and records in the IIS.

² The remaining 19% did not know the extent to which their IIS was impacted by incomplete data and entered a response of “unknown.”





Incomplete patient populations and vaccination histories in the IIS compromise the IIS's ability to provide accurate clinical decision support, inform population-based coverage assessments, properly support disease outbreaks, and drive policy and resource allocation decisions. This lack of completeness can be attributed to one or more of the following:

- Low provider participation
- Incomplete reporting of data:
 - Poor manual reporting from providers using direct data entry
 - Existing data in an EHR not being reported to the IIS once an electronic interface has been established
 - Gaps or inconsistencies in ongoing data feeds not being identified and/or properly resolved
- Patients simply not receiving all recommended vaccinations

The American Immunization Registry Association (AIRA) leveraged a community-informed approach to examine current methodologies for the collection of legacy data, as well as data gap monitoring and resolution practices. **Routinely pursuing legacy data and actively monitoring production interfaces are key strategies for ensuring that the IIS represents the most comprehensive picture of each individual's full immunization record, as well as the general vaccination status of the jurisdiction as a whole.** The goal of this project was to develop best-practice guidance for standardizing and improving processes for data backloads to increase the amount of complete patient and vaccination data available in IIS.

This project expands on the recently published AIRA resource *Onboarding Consensus-Based Recommendations*³ (November 2018) and builds upon the many resources that currently guide and facilitate IIS implementation best practices.

³ https://repository.immregistries.org/files/resources/5c377a4b2a490/aira_onboarding_recommendations_final.pdf

1.2 AUDIENCE, METHODOLOGY, AND DOCUMENT ORGANIZATION

This document is intended for practical use by IIS program staff, EHR vendors, and immunization providers. The guidance in this document is meant to encourage IIS programs to work with their providers and EHR vendor partners to improve IIS data saturation to promote better clinical decision support and public health assessment, outreach, and response. Information for this project was gathered through an AIRA survey, interviews with select IIS programs and EHR vendors, and review of existing documents and resources.

The survey was used to gather general information about how and when IIS programs typically collect legacy data, common challenges and barriers around the collection of this data, and best practices for identifying and resolving gaps in IIS reporting. The survey resulted in 35 completed surveys, each representing a unique jurisdiction. The information from the survey was used to identify subject matter experts (SMEs) for the interview portion and to seed the interview question set for richer discussion.

Eleven interviews were conducted in conjunction with this project, including nine IIS programs and two EHR product vendors. A complete list of interviewed participants has been included in [Appendix D. Acknowledgments](#). The information collected during the interview discussions was summarized and paired with the survey findings to develop the narrative and advise the implementation considerations and recommendations that appear in this document.

This document provides an overview of the most common import methodologies and timing strategies for pursuing legacy data, and it addresses community concerns about validating legacy data, data quality, and system performance considerations. The narrative also examines alternative data sources and challenges with provider participation. The final section offers best practices for the monitoring and resolution of missing patient and vaccination records (“gaps”). A list of implementation considerations and recommendations appears at the end of each topic section. These individual lists have been consolidated into a single resource in [Appendix B. Summary of Implementation Considerations and Recommendations](#).

A complete listing of the resources used to inform this document has been included in [Appendix C. References](#).

SECTION 2 GUIDANCE, CONSIDERATIONS, AND RECOMMENDATIONS

The following sections address the best methods and timing strategies for pursuing legacy data loads, as well as some of the most common challenges and implementation considerations.

IIS programs can leverage this information to develop plans for prioritizing and routinely securing legacy data in conjunction with onboarding activities and to revisit the possibility of securing legacy data for existing interfaces.

The following figure provides a sample workflow for how an IIS program might engage with an EHR vendor/provider for selecting legacy data and transmitting these records to the IIS.

Figure 1 | *Sample legacy load process*



Establish an agreed upon format/method

Establish a general timeline

Address common data quality concerns

Address barriers and challenges

**Schedule the date and time for
executing the load**

**Establish a plan for routine monitoring
and gap resolution**

2.1 LEGACY IMPORT METHODOLOGIES

The best method for a successful legacy data transfer ultimately depends on the EHR system's capability for extracting and transmitting the requested records. **IIS should offer and support a variety of options in order to maximize the likelihood of securing legacy data across disparate EHR systems.** Some EHRs may charge a fee to produce a custom data extract. These fees cover the man-hours to design and develop the extract, verify the content, and make any IIS-requested adjustments.

In general, **most EHR vendors will prefer to leverage the established EHR-IIS HL7 interface to transmit legacy data.** This may be accomplished using either an HL7 single-patient protocol or an HL7 batch protocol (multi-patient). A traditional, delimited flat file methodology is often reserved for specific corner cases and is generally being phased out across the IIS community for other routine data exchange. Direct data entry is always an option regardless of electronic data exchange, but it might not be a feasible expectation for most mid- to high-volume providers.

Common legacy import methods:

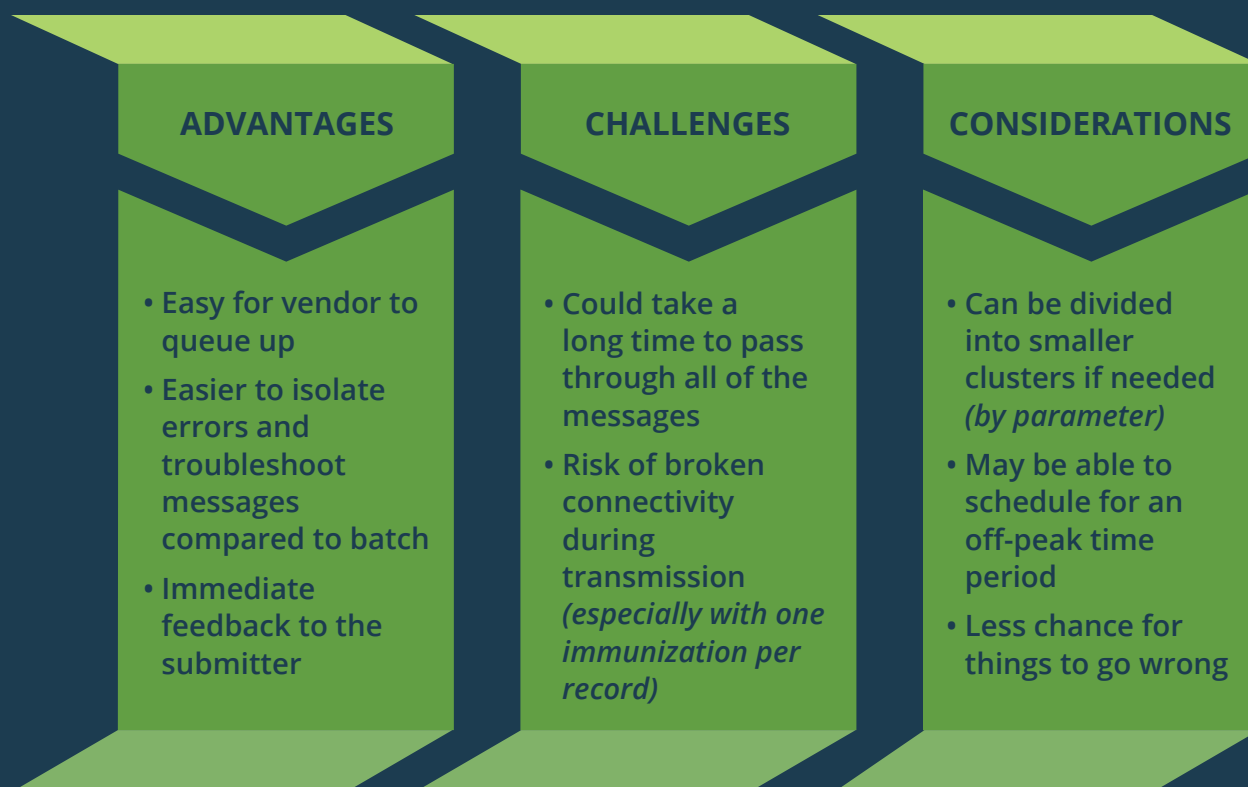
- HL7 (single-patient)
- HL7 batch (multi-patient)
- Traditional flat file (delimited)
- Direct data entry

The following narrative describes the various methods for transmitting legacy data, along with the advantages, challenges, and considerations for each approach.

2.1.1 HL7 (SINGLE-PATIENT PROTOCOL)

This method involves using the HL7 interface to facilitate submission of legacy messages from an EHR to an IIS, one patient at a time. This may include one patient-one vaccination per message or one patient-multiple vaccinations per message. Legacy data submitted using single-patient protocol can be facilitated through several different approaches:

1. EHR vendor identifies ("flags") all patient and vaccination records not previously reported to the IIS and queues the records to flow through the HL7 interface in a series of back-to-back messages.
2. IIS program develops a procedure to load data from an **HL7 batch file** using single-patient protocol (refer to following section).
3. EHR includes a patient's full vaccination history with each new VXU submission.

Figure 2 | HL7 single-patient assessment

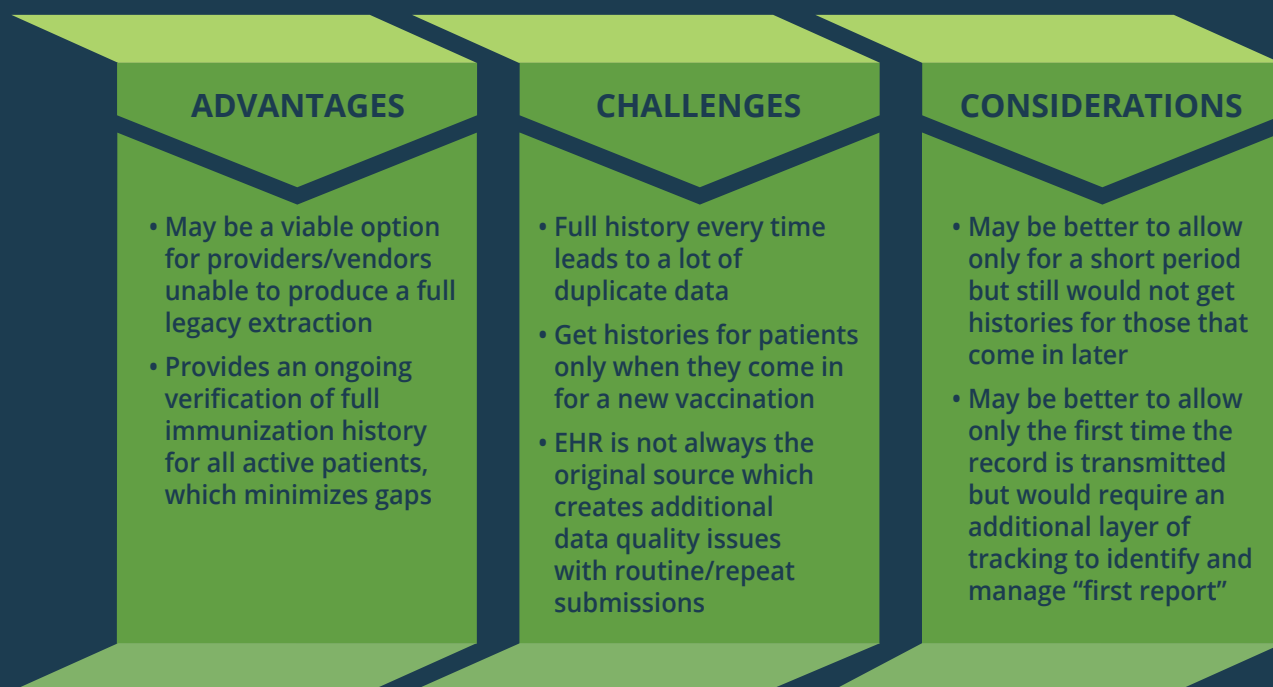
Note: IIS should not assume that single-message submission is necessarily slower and less efficient than a batch submission. HL7 single-patient protocol is not inherently slower at inserting data into an IIS than sending data using a comparable batch protocol. The speed of processing depends on factors in both the sender and receiver, and it is technically possible for rapid single-message submission to operate faster than a batch submission.

FULL VACCINATION HISTORY WITH NEW VXU (AS A METHOD) – SUBCATEGORY OF HL7 SINGLE-PATIENT PROTOCOL

Full vaccination history with a new VXU message is both an HL7 single-patient method and a timing strategy. As a method, full history with VXU leverages the EHR-IIS HL7 interface to submit a patient's full immunization history each time a new VXU submission is initiated. While most IIS offer support for this method, it is generally not a preferred strategy due to the level of deduplication needed to support it. With this method, legacy data is transmitted only when the patient is seen for a vaccination encounter. As a result, legacy data will trickle in over time, but the full history will be transmitted every time the patient receives a vaccination resulting in duplicate reports of previously deduplicated records. This creates the potential for additional data quality issues. This option is worth exploring, however, with providers that are unable to produce a one-time legacy data extract.⁴

⁴ Under the ONC Notice of Proposed Rulemaking (NPRM) to Improve the Interoperability of Health Information, DHHS has proposed improvements for the access, exchange, and use of electronic health information (EHI). The proposed rule focuses on exporting data from one EHR to another when a provider changes its health IT system; however, IIS may be able to leverage this same export process for securing legacy vaccination data from an EHR. IIS programs should monitor this rule as discussions and activities regarding this issue further evolve. <https://www.healthit.gov/topic/laws-regulation-and-policy/notice-proposed-rulemaking-improve-interoperability-health>

Figure 3 | *HL7 single-patient assessment*





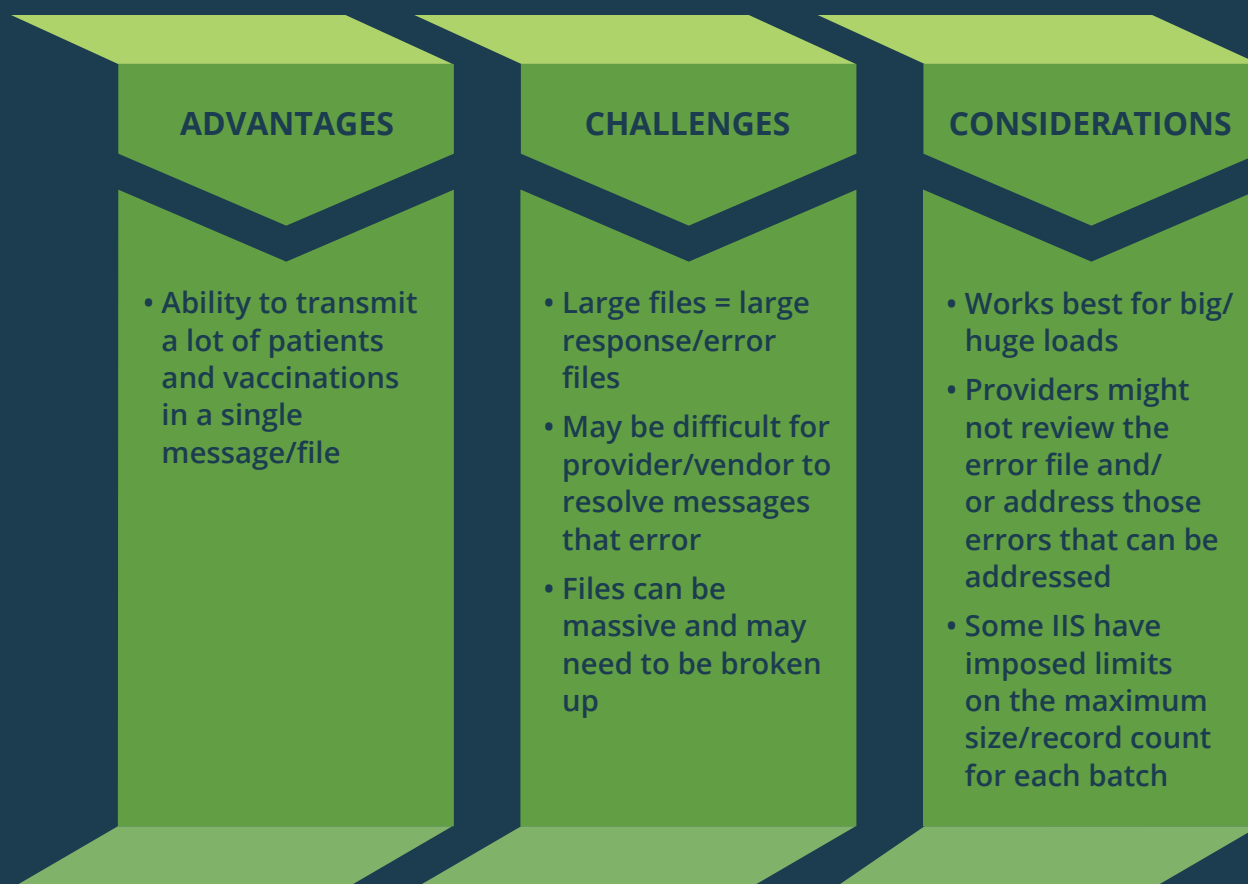
2.1.2 HL7 BATCH (MULTI-PATIENT PROTOCOL)

The term “HL7 batch” is used inconsistently across both the IIS and EHR communities. Essentially, there are three primary definitions being applied interchangeably for this term:

1. **True batch** – Uses the HL7 batch file specification as defined in the *HL7 Version 2.5.1 Implementation Guide for Immunization Messaging*⁵ (batch header/trailer and file header/trailer). This process leverages a direct transmission of the batch message(s) from the EHR to the IIS using an established EHR-IIS HL7 interface.
2. **Queued batch** – Vendor/provider uses a process to identify and queue a “batch” of records and then submit them using the [single-patient protocol](#) (refer to previous section).
3. **HL7 message file (sometimes referred to as an “HL7 flat file”)** – Employs a process where numerous individual HL7 messages are saved to a file/document that can be posted to an SFTP folder and loaded to the IIS at some later point using either an [HL7 single-patient](#) or true-batch protocol. This is an indirect process that leverages an SFTP as an intermediate step in transmitting a batch of HL7 messages from the EHR to the IIS.

For the purposes of this narrative, HL7 batch will be used to describe any method that results in multiple patients and vaccinations being submitted in a single-batch submission or series of batched messages. **HL7 batch is generally the preferred method for transmitting large quantities of legacy data, with true batch and queued batch as equally popular strategies.**

⁵ <https://www.cdc.gov/vaccines/programs/iis/technical-guidance/downloads/hl7guide-1-5-2014-11.pdf>

Figure 4 | *HL7 batch assessment*

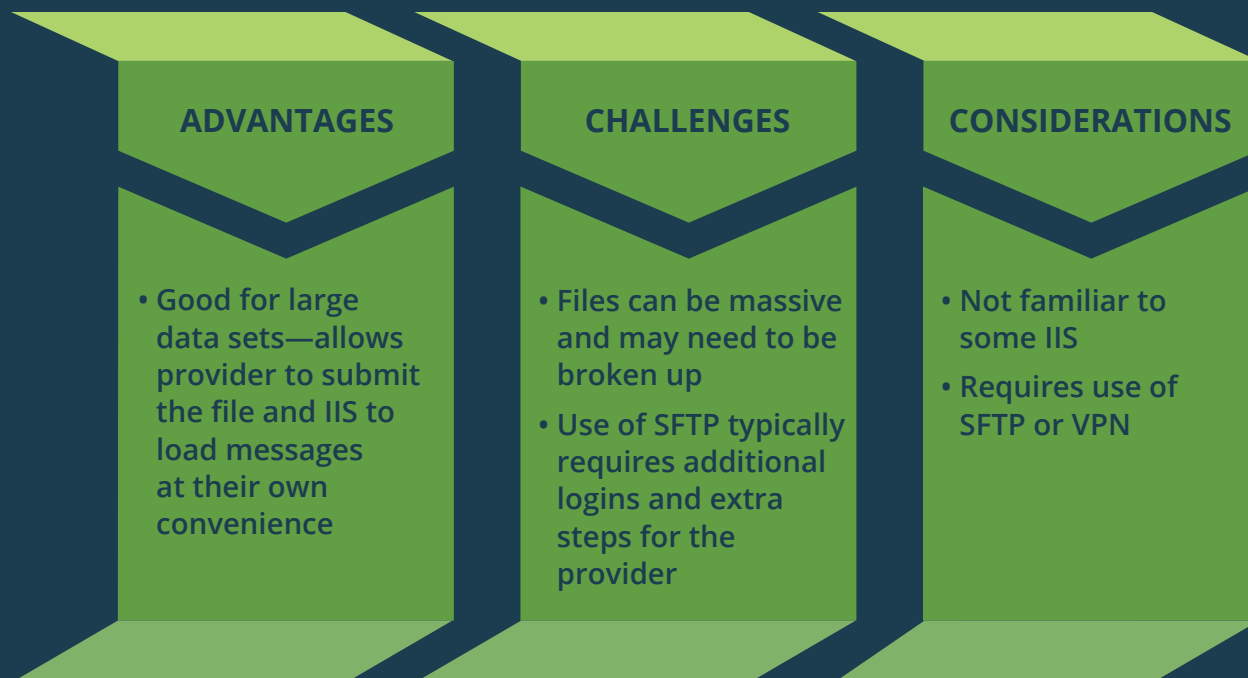
HL7 MESSAGE FILE OR “HL7 FLAT FILE” (SUBCATEGORY OF HL7 BATCH)

As noted under both the HL7 single-patient and HL7 batch message protocols, there is an option to leverage an HL7 message file or “HL7 flat file.” This method leverages a flat document or file that contains multiple HL7 messages. The document or file is then loaded to a secure location using SFTP or VPN. Loading of these records into the IIS would then be handled manually or through an automated loading procedure. The advantage of this option is that the IIS can review and load the records from the document/file at their convenience.

This method is currently preferred by one of the larger EHR partners because it provides an easy process for them to compile and transmit significant amounts of patient and vaccination data. This methodology is not overly common, however, and may likely be considered a custom extraction effort by most EHR partners.

There is another possible use case for leveraging HL7 message files in the IIS community. For instance, one IIS program has developed an automated process for sending HL7 messages to a file during system upgrades, scheduled and unscheduled downtime, and unforeseen situations where messages may not be processed in a timely manner. In these situations, the IIS program will then retrieve and load the messages at a later time. This ensures that all messages are securely captured and ultimately uploaded to the IIS to mitigate incomplete data loads or other potential data gaps. Data gaps are discussed further in the sections [Gap monitoring](#) and [Gap resolution](#).

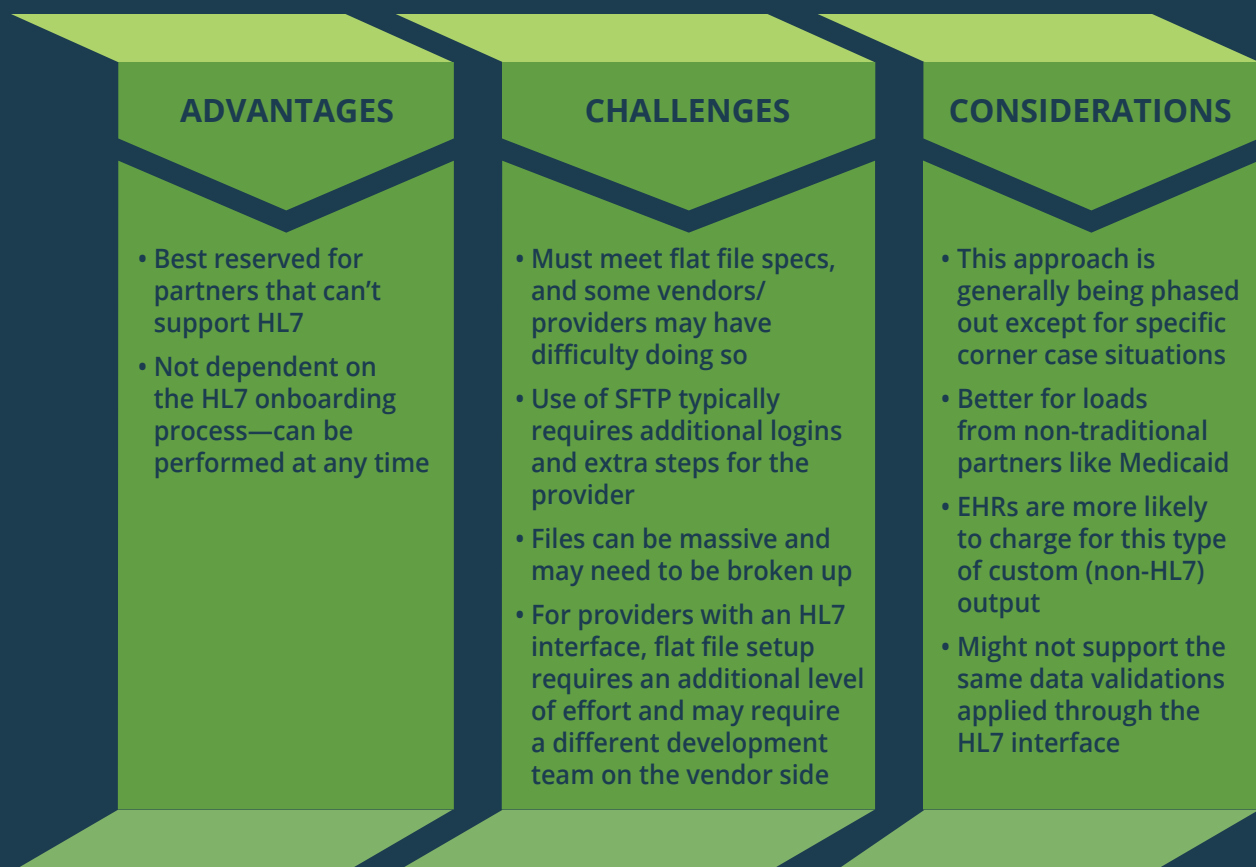
Figure 5 | *HL7 message file “HL7 flat file” assessment*



2.1.3 DELIMITED FLAT FILE

A traditional, delimited flat file does not leverage HL7. This method requires the provider or vendor to extract legacy data into a file conforming to an IIS flat file specification using a specified delimiter. The file is then uploaded directly into the IIS interface or posted to a staging folder using SFTP to be loaded later by IIS staff or an automated protocol. Delimited flat files are considered to be an older technology and are primarily reserved for providers without HL7 capabilities, very small providers, and other data partners, like Medicaid and Vital Records. There may be other use cases where a traditional flat file option may be appropriate and/or necessary in order for the IIS to obtain legacy data. These situations should be assessed on a case-by-case basis.

Figure 6 | *Delimited flat file assessment*



2.1.4 DIRECT DATA ENTRY

This option is available to any user with access to the IIS and appropriate user permissions. This method does not leverage HL7 messaging and can be performed without regard to the onboarding process. This method is typically more labor intensive, especially for providers with a large patient population. Manual entry also introduces an increased potential for data entry errors or inconsistencies in reporting practices. While certainly a viable option when other alternatives are not available, the decision to leverage direct data entry should be assessed from a cost-benefit perspective.

Figure 7 | *Direct data entry assessment*

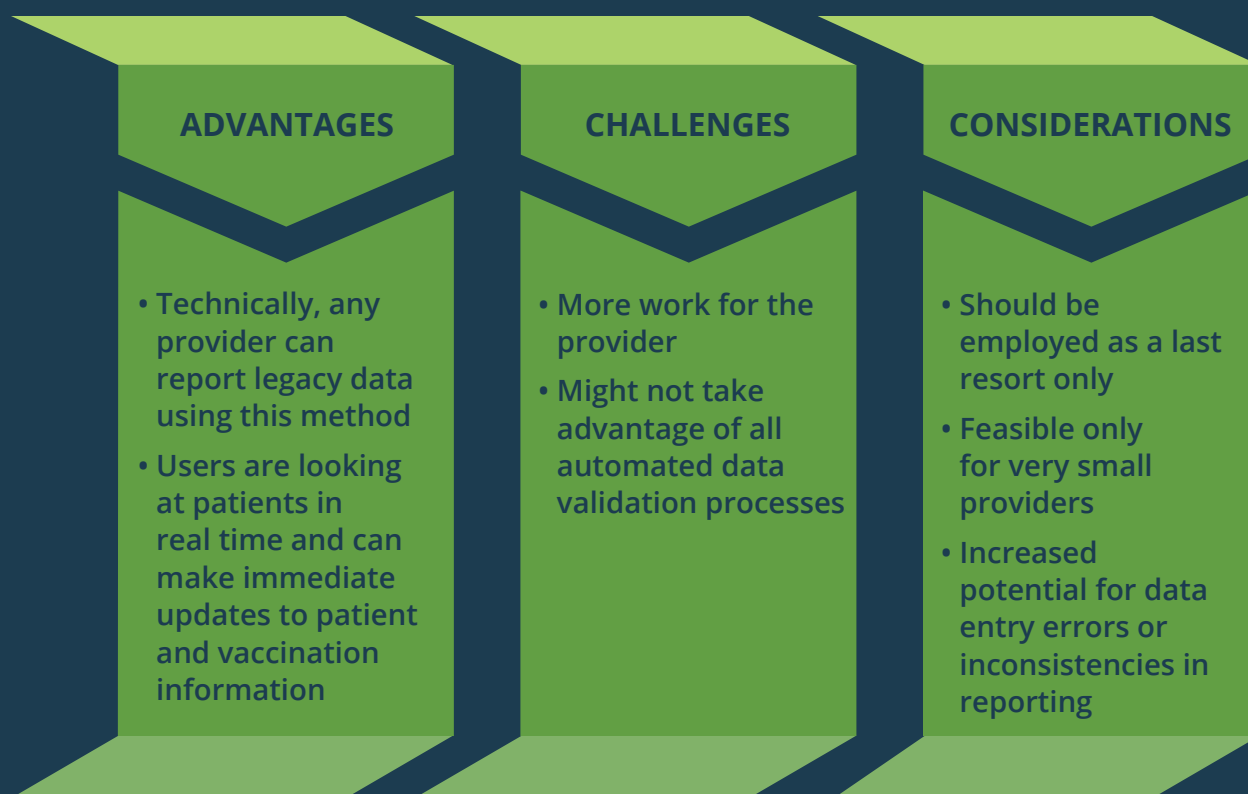
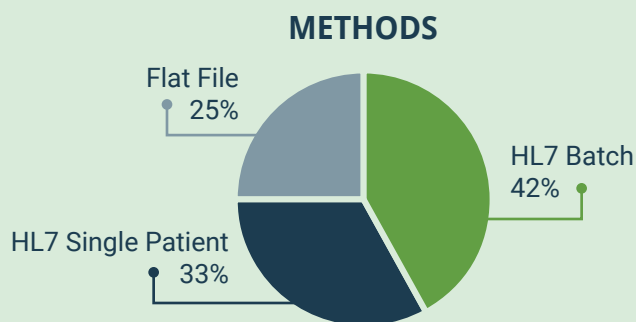




Figure 8 | Methodology findings from interviews and community survey

Interview participants echoed community survey responses (n=35) noting HL7 batch (multi-patient) as the preferred protocol for legacy data, followed by HL7 single-patient and then traditional delimited flat file.



Implementation Considerations and Recommendations: Methodologies

- Offer a variety of methods to support a range of EHR capabilities and preferences.
- Leverage the IIS-EHR HL7 interface for legacy loads unless there are other prevailing circumstances.

2.2 TIMING OF LEGACY IMPORTS

Mature IIS may already have a high degree of data saturation from long-standing provider interfaces or direct data entry requirements backed by mandatory reporting laws. For these systems, the pursuit of legacy loads might not be a program priority, and the timing of any such pursuits is not as critical as it may be with newer IIS or those systems with a lesser degree of data saturation.

In general, it is best to secure legacy data in conjunction with HL7 onboarding activities. To improve the chances of facilitating an exchange of legacy data, **IIS programs should set the expectation for legacy transmission early in the onboarding process**, preferably as a discussion topic during the initial kick-off call with the provider and its EHR vendor. This discussion should address the provider's/vendor's capabilities for producing a legacy extract and the preferred method/format for doing so.

Once the method/format has been determined, stakeholders should agree on the timing for this transmission. **There is general consensus that the best time to secure legacy data is immediately after an interface is approved and moved to production.** Technically, a legacy load could be transmitted during active onboarding as soon as the message validation phase is complete, but most stakeholders prefer to wait until production approval. If an IIS program tries to pursue legacy data at a later point, it may be significantly more difficult because the provider/EHR vendor onboarding teams are no longer actively engaged. Methods such as direct data entry or traditional flat file could be performed at any time because these methods have no dependencies on the establishment of an HL7 interface.



Table 1 | *Considerations for each of the primary timing options*

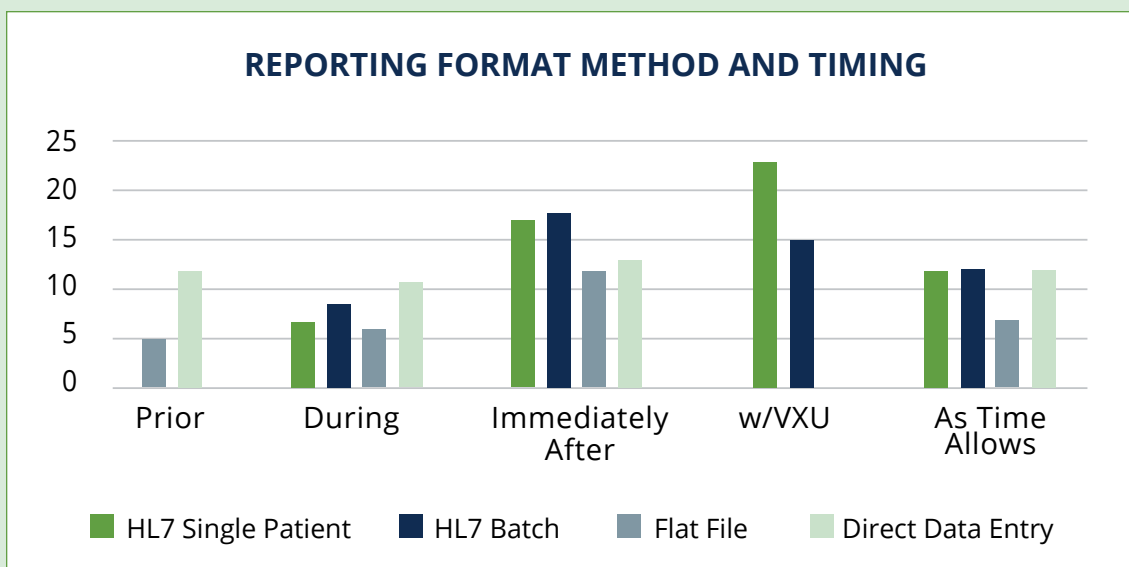
LEGACY TIMING	CONSIDERATIONS
Prior to onboarding	<ul style="list-style-type: none"> No way to know how much of the existing data has been reported to the IIS Direct entry and flat file are the only viable formats
During onboarding	<ul style="list-style-type: none"> Not a viable option until the HL7 interface has been tested and validated Immediately following message validation portion may be an option, especially if the IIS wants to test legacy data in QA first
Immediately after onboarding	<ul style="list-style-type: none"> Provider and vendor are still actively engaged Best chance of getting a legacy load during this time period Easiest for vendor because of ramped up staffing May be best to load legacy data first before opening the gates to new message submissions
Full vaccination history with new VXU (as a timing strategy)	<ul style="list-style-type: none"> May be a viable option for providers/vendors unable to produce a full legacy extraction Gets histories for patients only when they come in for a new vaccination
As time allows	<ul style="list-style-type: none"> Harder to pursue a legacy load for existing interfaces

Note: While the period immediately following onboarding is the optimal time for securing a legacy data load, IIS programs with low data saturation should continue to explore opportunities to get data from providers with existing IIS-EHR interfaces. One possible approach is for the IIS to consider issuing a communication to providers with older interfaces that can serve multiple purposes: (1) remind providers of the ongoing roles/responsibilities for interface monitoring and maintenance, (2) request updated contact info, and (3) request a legacy data load for any doses that predate their interface with the IIS.



Figure 9 | Community survey findings on reporting format by method and timing

Community survey respondents (n=35) reported the most common import methodologies based on general timing of legacy uploads to the IIS. This chart reflects the results of these responses.



Implementation Considerations and Recommendations: Timing

- Discuss the legacy data request with each provider early in the onboarding process, preferably during the initial kick-off call.
- Prioritize legacy data loads for the period immediately after onboarding.
- Develop a strategy to secure legacy data from already existing provider interfaces.

2.3 VALIDATION OF LEGACY RECORDS

Legacy data loads are typically subjected to the same validation processes and protocols as routine data submissions. **Legacy messages should conform to HL7 technical validations (message format, required fields “R,” and conforming code sets) but may have different thresholds for general data quality due to changes in codes, requirements, and IIS expectations over time.** For instance, IIS should expect to see previously active CVX codes, historically “administered” doses that don’t include all currently required data elements, and missing VFC segments for funding source and dose-level eligibility. With legacy data, two things are true:

1

Data quality in legacy records is unlikely to meet current data quality standards, with diminishing quality observed with increasingly historic vaccination dates.

2

Providers may be unwilling to invest the time/resources to improve older records and, more importantly, unable to correct legacy data.

There are two prevailing philosophies around the collection of legacy data where some IIS programs believe that **“no data is better than incomplete or inaccurate data,”** while others believe that **“some data is better than no data.”**

The first philosophy of accepting only the highest-quality records ensures that IIS data quality standards are strictly maintained but also indicates that the IIS program is willing to forego the addition of records that can’t meet these higher standards. Patient and vaccination records that are incomplete or contain inaccuracies per current standards do not pass IIS data quality validations and are not added to the IIS database.

Figure 10 | *Differing philosophies*



With legacy data, most IIS programs tend to favor the second philosophy and will accept whatever records the IIS can acquire by focusing on only the most critical data elements (e.g., vaccine type and vaccination date). This allows the IIS to fill in data gaps and round out patient records that are missing one or more vaccination events. With these core data elements, patient records become more complete, vaccine forecasting is improved, duplicate doses are prevented, and missed opportunities to vaccinate are reduced. Some data elements such as lot number and VFC eligibility status are important for point-in-time accounting of active inventory but become increasingly less relevant for records that are several months or years old. Incomplete records or those that contain inaccuracies are added to the IIS database but can be identified and assessed as needed using existing IIS data quality tools.

The philosophy adopted by an IIS program may ultimately depend on the maturity of the IIS and/or perceived level of current data saturation levels. **From a provider/EHR vendor perspective, however, if the resources have been invested to produce the legacy load, the IIS should attempt to consume any and all viable records.** Some IIS programs may argue that the provider/EHR should not submit records that won't pass current data quality standards. This approach places additional burden on the provider/EHR vendor to determine which records should and should not be transmitted. A better practice is to allow the IIS business rules for processing records to determine which records should be accepted or rejected.

Knowing that data quality in legacy loads may be inferior, most IIS mitigate concerns about data quality by coordinating a specific time with the provider to execute the legacy load. This allows the IIS to prepare for the increased errors and warnings that may result and isolate the issues associated with the legacy data load as a separate activity from the provider's routine submissions. When the legacy load is isolated, it allows the IIS program to strategically leverage its data quality assessment tools, if desired, to identify and target issues that could potentially be resolved and resubmitted. The section [Data quality concerns](#) will address the most common challenges encountered with legacy data loads.

The following list represents some additional suggestions offered by the interviewed subject matter experts that may help with processing legacy data.

- **Apply an error threshold.** For instance, if the error rate (rejection or severe errors) is greater than 50%, reject the entire file and work with the provider/vendor to resolve issues if possible and resubmit.
- **Assess the legacy load in QA first before running it into the production environment.** This activity has the benefit of allowing IIS staff to preview the data before it gets submitted to production; however, this adds an extra step and may incur additional resources with arguable gains when legacy data is known to be of lesser quality and errors are unlikely to be addressed. This approach may be most beneficial in situations where there are significant data quality concerns for a particular provider. An IIS program may also consider running only a portion of the legacy load to get a feel for how the data will behave without performing a test load in its entirety.
- **Focus on more recent data submissions.** Some IIS programs have conducted studies that demonstrate diminishing returns in data quality as the dates of administration get further back in history. IIS programs may decide to portion out the legacy data based on vaccination date and load in phases. For example, the first phase would include all records with vaccination dates in the past five years; the second phase would include vaccination dates from the past 6 to 10 years. As the number of errors increases beyond an acceptable threshold, the IIS could decide to terminate additional loads of older data.

Implementation Considerations and Recommendations: Validation

- Apply the same data validation methods and criteria used for routine data submission (e.g., HL7 technical validations).
- Accept as much viable data as possible even if legacy records do not conform to current data quality standards.
- Schedule the timing of the legacy load with the provider/EHR vendor so IIS program staff can account for an increased spike in errors and warnings.

2.3.1 DATA QUALITY CONCERNS

The most common data quality issues observed with legacy data imports include:

- Duplicate patient and vaccination records
- Historically “administered” doses that do not include all currently required fields
- CVX code management and unspecified vaccine types
- Generic vaccination dates and questionable data sources

2.3.1.1 Legacy data creates duplicate patient and vaccination records

One of the biggest challenges with importing legacy data into an IIS is the introduction of duplicate data for both patients and vaccinations. While the primary motivation of pursuing legacy data is the collection of previously unreported patient and vaccination records, legacy data loads will likely include many records previously known to the IIS. Good patient-level and vaccine-level deduplication logic is critical for the proper processing of legacy data. When deduplication algorithms have tuning issues that result in bad matches or improperly resolved records, it results in a significant allocation of IIS staffing resources to manually review and resolve duplicates.

IIS programs should periodically assess their patient and vaccination matching logic and deduplication strategies to ensure that these IIS functions produce the expected results and outcomes. Over time deduplication algorithms may need to be retuned to support evolving data trends and IIS functionality. **Deduplication procedures should be documented to include the various rules being applied to data matching and duplicate resolution.** This documentation should also include a listing of which fields are most important in the deduplication process and which situations the IIS process will be unable to handle.

Though many IIS programs complain about manually reviewing and resolving duplicate records, most programs prefer a more conservative matching approach. While a more conservative approach results in a higher number of manual reviews, it reduces the risk of false merges that may result from more aggressive automated match strategies. In cases where an inappropriate merge may occur, the IIS should be equipped with the tools and ability to separate improperly merged records.

Implementation Considerations and Recommendations: Duplicates

- A good patient-level and vaccine-level deduplication algorithm for identifying and resolving duplicates is critical for the proper processing of legacy data.

2.3.1.1.1 Duplicate Patients

The goal of patient deduplication is to compare core fields such as unique patient ID, name, date of birth, birth order for multiple births, sex/gender, and demographic variables to identify possible matched records and then resolve duplicates by automating the merging of patients or queuing them for manual review. **When duplicates exist in the IIS, it compromises the integrity of the data, affects the ability of the IIS to offer appropriate clinical decision support, and artificially inflates denominators for community and provider-level coverage assessment activities.**

There are two primary documents that provide guidance to IIS programs for the implementation of patient-level deduplication:

- *Immunization Information Systems Patient-Level De-duplication Best Practices* (CDC, 2013)⁶
- *Consolidating Demographic Records and Vaccination Event Records* (MIROW, 2017)⁷

The 2013 CDC *IIS Patient-Level De-duplication Best Practices* guidance document is a foundational reference for IIS deduplication development and replaces a previously published CDC guide from 2002. The 2013 guide reexamines IIS-based patient deduplication and provides updated best practices for streamlining, standardizing, and improving the patient deduplication process. The CDC guidance addresses the methodologies for match candidate selection and scoring algorithms to determine which records are automatically deduplicated and which are queued for manual review.

⁶ <https://www.cdc.gov/vaccines/programs/iis/interop-proj/downloads/de-duplication.pdf>

⁷ <https://repository.immregistries.org/resource/consolidating-demographic-records-and-vaccination-event-records/>



Due to the historic nature of legacy data loads, it may prove particularly problematic to correct name misspellings, data entry errors for date of birth, and differences in demographic details that have changed over time. The issue becomes even more complicated with the addition of consent requirements, adoptions, name changes, and record deletions. The 2013 CDC guidance document offers a few best practices that may be particularly relevant for addressing these issues.

- **Standardize first and last names:** The IIS should remove hyphens to assess as a two-word name, remove punctuation, assess as all caps or caps-neutral, and consider implementing phonetics and common nicknames.
- **Standardize addresses:** The IIS should investigate the incorporation of the SmartyStreets⁸ address service offered through AIRA or a similar USPS address standardization tool. If this is not feasible for the jurisdiction, consider breaking the address into smaller components (house number, street name, street suffix, etc.) and eliminating formatting filler such as hyphens and parentheses in phone numbers.
- **Use a combination of deterministic and probabilistic algorithmic methods:** By combining the two strategies, IIS can perform more sophisticated analysis and build on the strengths of both approaches.
- **Increase automation:** Leverage newer technologies to promote machine learning and improve patient matching and resolution to decrease manual staff review.
- **Recognize records that have been previously resolved:** Identify methods to cache previously resolved records so the IIS/IIS staff don't have to repeat the deduplication process each time the pairing is identified.

⁸ <https://www.immregistries.org/address-cleansing>



The 2017 MIROW *Consolidating Demographic Records and Vaccination Event Records*⁹ guidance document focuses on the consolidation of a matched record set by selecting the best variables from each record to produce a better, more complete final record. These activities apply to demographic records as well as vaccination records and vaccination events (see also duplicate vaccinations discussion to follow). The MIROW guide offers a series of principles and business rules for defining what constitutes “best” when assessing each data element in a pair of demographic records.

Note: In general, the MIROW guidelines favor the record with a more recent submission date; however, with legacy data loads, the submission date will be misleading, and IIS may want to fall back on other criteria to ensure that more recent data do not get overwritten by a legacy update.

When determining which demographic record or record components to retain, the 2017 MIROW guide suggests the following:

- **Select the best data-element-level characteristics:** This can be determined by assessing for valid data values and looking at which data are more complete and/or more specific. The IIS should then use the “best” value for each data element.
- **Select the best record-level characteristics:** This can be determined by assigning a confidence level to each record based on submission method, submission type, data source, and submission date.

⁹ <https://repository.immregistries.org/resource/consolidating-demographic-records-and-vaccination-event-records/>

2.3.1.1.2 Duplicate Vaccinations

There are two primary documents that provide guidance to IIS programs for the implementation of vaccine deduplication logic:

- *Vaccination Level Deduplication in IIS* (MIROW, 2006)¹⁰
- *Consolidating Demographic Records and Vaccination Event Records* (MIROW, 2017)¹¹

The primary guidance for vaccine-level deduplication was defined in the 2006 MIROW guide. This document provides principles and business rules for the identification (evaluation phase) and resolution (resolution phase) of duplicate vaccination events. **The MIROW guidance applies to all administered-historical, administered-administered, and historical-historical vaccination event comparisons.**

The 2017 MIROW guidance on consolidation replaces Chapter 5: Resolution Phase in the 2006 MIROW deduplication guide and recommends that IIS work toward the use of a single consolidated vaccination event to best support all IIS functions, including clinical decision support and query/response. Most IIS are still applying the “best record” methodology from the 2006 guide but may be at various stages of implementing the new 2017 guidance. **The advantage of a consolidated vaccination event record is that it accounts for the “best” variables presented in two matched records and results in a “better” record than either record on its own.**

What determines “best”? Regardless of whether an IIS applies a sequential (deterministic) or weights-based (probabilistic) approach for resolving duplicate vaccination events, the best record or the best variable can generally be determined by evaluating the following elements:

- **Administered-historical indicator:** Identifies the record source type and whether the vaccination event is a “historical” or “administered” report. In an administered-historical comparison, the administered vaccination typically takes precedence, whereas an administered-administered or historical-historical comparison will require a more in-depth analysis of other data elements.
- **Submission method:** This defines whether the information was entered directly through the IIS user interface or through an electronic interface. Some IIS may trust data entered via the user interface more than electronically submitted data.

¹⁰ <https://repository.immregistries.org/resource/vaccination-level-deduplication-in-immunization-information-systems-1/>

¹¹ <https://repository.immregistries.org/resource/consolidating-demographic-records-and-vaccination-event-records/>

- **Completeness:** Records that are more complete based on the population of “important” variables will typically take precedence. MIROW suggests that the hierarchy of important variables is defined as follows:
 - Level 1: Determining a confidence level—who submitted the record, method used to report the record to the IIS, record source type (administered-historical), and documentation type (clinical, billing, or transcribed)
 - Level 2: Presence of a lot number or recorded as a combination vaccine
 - Level 3: Vaccine type (CVX) and/or presence of a trade name
- **Specificity:** Records with more specific information will typically take precedence over less specific values, and populated values (if valid) would take precedence over a previously null value. For instance, a specific vaccine type is better than an “unspecified” report, and a combination vaccine is typically preferred over individually reported vaccine components of the combo. Further, a populated trade name or manufacturer is typically preferred to a record with a null or less specific value.

Records reported through a legacy load may present other unique challenges when compared to routine VXU submissions. For instance, the administered-historical indicator might not be reliable, especially for programs that opt to flag all doses in a legacy load as “historical.” Reporting of single antigens versus a combination vaccine may be complicated by preexisting EHR functionality that determined how a historically administered combination vaccine was originally recorded. Generic dates (e.g., 1/1/YYYY or 15th day of the month) and original source (e.g., parent report or old paper records) may make it more difficult to identify match candidates or verify/resolve older data. See additional discussion on these topics to follow.

2.3.1.2 Historically “administered” doses do not include all currently required fields

Due to changes in data entry requirements over time, historically “administered” doses might not contain all fields requested under the current standard¹² (e.g., lot number, manufacturer, expiration date, dose-level eligibility, or funding source). With more stringent HL7 validations, some IIS may fail doses reported as administered if they do not contain all required fields. Other IIS may accept the viable portions of the message but send back a warning in the acknowledgement (ACK) that required fields are missing. As noted previously, it is unlikely that a provider would be able to locate missing information or invest the time and resources necessary to update legacy records.

¹² <https://www.cdc.gov/vaccines/programs/iis/technical-guidance/hl7.html>

As recommended in a previous section, IIS should leverage the same HL7 technical validations applied to routine data loads, but SMEs interviewed for this project suggest that IIS also consider the following approaches:

- **Do not reject historically “administered” doses even if currently required fields are missing.** In legacy data, vaccine type and vaccination date are the most critical data elements needed to complete a vaccination record and impact forecasting. The IIS should import the vaccination event “as is” and retain all viable portions. If IIS logic is too stringent, consider setting all legacy doses to import as “historical” (see below).
- **Assess the legacy data prior to loading to production.** This may include a pre-test in the QA environment, as previously noted, or a discussion with the provider about “administered” data quality during the planning phase. If a test run produces an unacceptable number of errors or if a provider expresses concern about the data’s ability to pass the IIS’s data validation process for administered doses, consider setting all legacy doses to “historical” (see below).
- **Report all legacy doses to the IIS as “historical” regardless of how they were recorded in the EHR.** This would allow messages recorded as “administered” in the EHR to bypass more stringent IIS requirements for the reporting of current dose administrations. *Note: this option may involve record modification and may present additional considerations concerning vaccine deduplication.* If this is not an option, consider using an alternative methodology for loading legacy data.



Implementation Considerations and Recommendations: Historically “Administered”

- Use standard HL7 technical validations for accepting/rejecting records, but if legacy messages flagged as “administered” are likely to fail, consider using historical flags or an alternative method to load (e.g., traditional flat file).



2.3.1.3 Legacy data includes inactive CVX codes and unspecified vaccine types

There are several challenges related to CVX codes in legacy data. The following table includes the most common issues and suggested solutions.

Table 2 | Common CVX code challenges and solutions

ISSUE	SUGGESTED SOLUTION
<ul style="list-style-type: none"> Legacy data contain CVX codes that are no longer active. 	<ul style="list-style-type: none"> Support a master CVX code table. Use date of administration to determine whether the CVX code was active at the time.¹³
<ul style="list-style-type: none"> Legacy data contain a large volume of unspecified vaccine types (NOS). 	<ul style="list-style-type: none"> Allow NOS vaccine types on historical doses. When possible use master CVX code table to derive a more accurate vaccine type.
<ul style="list-style-type: none"> Legacy data may include increased incidence of cross-coding (e.g., PCV-7, PCV-13, and PPSV-23). 	<ul style="list-style-type: none"> Routinely leverage data-quality monitoring tools to identify possible cross-coding based on age of patient, date of administration, and gender. Apply CDSi logic to determine what is/is not appropriate.
<ul style="list-style-type: none"> Legacy data includes CPT codes. CPT codes are not always a one-to-one match with CVX codes. 	<ul style="list-style-type: none"> Utilize resources on the CDC website to map CPT codes to acceptable CVX codes if using the HL7 interface to load legacy files.



The following resources are available on the CDC website to help IIS programs navigate common CVX code challenges:

- *IIS: Current HL7 Standard Code Set – CVX – Vaccines Administered*¹⁴
- *CPT Codes Mapped to CVX Codes*¹⁵
- *Clinical Decision Support for Immunization (CDSi)*¹⁶

¹³ Existing resources for determining “active” dates for CVX codes are somewhat limited. The current code tables do not reflect start and end dates. Dates that are presented represent a “last updated date,” which might not reflect the actual date that a particular code became active/inactive. There is an ongoing community need to formalize start and end dates for each CVX code.

¹⁴ <https://www2a.cdc.gov/vaccines/iis/iisStandards/vaccines.asp?rpt=cvx>

¹⁵ <https://www2a.cdc.gov/vaccines/iis/iisStandards/vaccines.asp?rpt=cpt>

¹⁶ <https://www.cdc.gov/vaccines/programs/iis/cdsi.html>

In some cases, the IIS business logic can be developed to automatically process or resolve these common issues: CVX codes that are no longer active, unspecified vaccine types that can be mapped to a more specific code, and translating CPT codes to CVX. In other cases, the records may require human intervention, such as cross-coding or pre-mapping of CPT to CVX, to achieve resolution. **For problems that affect more than a handful of records, the provider/IIS program could ask the EHR vendor to run a script to correct the problem records and resubmit, or the IIS can run a script against the IIS database to correct the problem records after they have been imported.**

Note: An increasing number of IIS and EHR vendors are offering support for the collection and transmission of NDC numbers. Most IIS will automatically map an NDC to its corresponding CVX code. While future legacy loads may include NDC numbers, it is unlikely that NDC will entirely replace CVX codes anytime in the foreseeable future. As an additional consideration, IIS do not currently support active/inactive dates for NDC numbers, and NDC numbers would be subject to the same challenges reported with inactive CVX numbers described above.



Implementation Considerations and Recommendations: CVX codes

- Assess CVX code validity based on date of vaccine administration by maintaining the CVX code table for all active and inactive CVX codes.

2.3.1.4 Legacy data includes generic vaccination dates and questionable data sources

On occasion legacy data may contain records where a generic date was used for the date of administration. The most commonly used generic dates include the 1st and 15th of a given month or 1/1/YYYY. The use of generic dates is seen most often with records from overseas, much older records, and influenza administrations. While generic dates may be difficult to identify, some SMEs reported that they have implemented business rules in their IIS to address the more common occurrences:

- Do not load any doses with 1/1/YYYY other than HepB, pneumonia, or Tdap.
- Code the 15th of the month into the deduplication algorithm. If a better record with a more accurate date comes in, select the record with the better date.
- Flag suspicious records and send to queue for manual review.

Another common concern expressed about legacy records is the reliability of the original data source (e.g., parent report). The original data source of a record is not often recorded in the EHR or reported to the IIS, so there is little opportunity to assess the source validity or develop business rules for record processing.

As a general rule for both generic dates and original record source, if the provider has recorded the dose as part of the patient's clinical record, the IIS should simply accept the dose at face value.

Implementation Considerations and Recommendations: Generic Dates and Data Sources

- Accept records with generic dates or questionable data sources at face value as reported by the provider.

2.4 PERFORMANCE

Due to the volume of records that may be transmitted during a legacy load, some IIS programs have expressed concern or have previously observed an impact on performance while legacy data is being received and processed. Performance impacts depend on a variety of factors, including the IIS architecture/resources, system configuration, and general bandwidth. For IIS programs that are concerned about performance issues, interviewed subject matter experts offered the following suggestions that may help with managing these concerns:

- **Schedule the timing of the legacy load** (date/time) and/or request a heads-up from the provider/EHR vendor before the transmission begins.
- **Schedule the load during off-peak periods** like evenings, early mornings, and weekends.
- Request that providers/EHR vendors **set a start/stop time** for periods of active transmission to prevent sessions from running into normal business hours.
- **Limit the size of each batch** (e.g., 2,000 records or 25MB) or set a maximum record count.
- **Break data into smaller clusters**, such as patient age groups (e.g., 0–6 years, 7–18 years, 18–30 years, etc.) or vaccination date ranges (e.g., past 5 years, past 6–10 years, etc.).
- **Monitor the load** to make sure it loads completely and does not crash the system.

In situations where the IIS is unable to keep up, the IIS may start dropping messages or sending back failed acknowledgements. In worst-case scenarios, it may even cause the system to crash. With dropped messages, the provider/EHR vendor might not know that messages were not received by the IIS or might not be able to pinpoint which records were/were not received, especially in large batch submissions. If the provider/EHR vendor receives errors/failures related to the transmission, the messages may get triggered again, thereby increasing the number of records hitting the IIS. **Queuing messages to an HL7 message file or batch file that can be loaded later by the IIS program is another strategy that IIS could consider for instances when IIS processing capacity is exceeded.**

Implementation Considerations and Recommendations: Performance

- Coordinate the timing of the legacy load with the provider/EHR vendor.
- Break up the transmission by a selected parameter, record count, or file size if needed.

2.4 PROGRAM PRIORITIES AND ELIMINATING BARRIERS

Some IIS programs do not routinely pursue legacy data from providers. There are a number of factors that drive this decision:

- IIS perception that data saturation has already been achieved. This is more common with mature databases backed by long-standing reporting mandates.
- IIS staff resources are limited, and legacy loads are not a priority or are facilitated only upon provider request.
- Concerns about data quality and risk of introducing poor quality data into the IIS.
- Consent laws that restrict the IIS from pursuing records without a consent flag or records dated prior to a certain point in time.

All IIS programs should ultimately strive to obtain complete immunization records for the entire population served by the IIS. This goal supports the core IIS mission of providing centralized records for use by all stakeholders, clinical decision support tools, and better population and provider-level coverage assessments. As more providers rely on their EHR platforms to support clinical workflows, the IIS should remain visible as the ultimate source for consolidated vaccination records and forecasting. HL7 bidirectional (query/response) interfaces provide an integrated solution for data sharing but might not be a viable option for all providers. In order to offer maximum value to all stakeholders and to best serve public health needs, providers should be able to access all known patients and vaccinations through the IIS. **IIS programs should continue to develop consistent messages and activities for promoting the value of an IIS and the importance of legacy data.**

Even for IIS programs that routinely prioritize the pursuit of legacy data, participation from the provider and EHR vendor communities may be a challenge. SMEs reported that many EHRs claim to be unable to send legacy data.¹⁷ This may or may not be true but could be a reflection of other factors, such as:

- The IIS offers only one or two methods for transmitting/importing legacy loads, and the vendor is unable to conform
- The IIS requirements for massaging, modifying, or updating legacy record data are too extensive for vendors to properly address
- The EHR vendor might not have much experience producing or executing a legacy file, especially if the feature isn't built into the existing code; the request for data is then treated as a custom effort

In addition to EHR capabilities, provider participation/investment is a related concern. Providers may be unwilling to authorize a legacy data extract if:

- The EHR vendor charges an additional fee to package and transmit the data
- The provider is unable or unwilling to invest resources (time and/or money) in improving legacy records at the request of the IIS
- The provider perceives that the majority of its data has already been reported to the IIS through a prior interface or direct data entry activities
- The provider implemented a new EHR system and did not migrate its old data to the new platform (in this scenario, the provider may have limited means for accessing its own legacy data)¹⁸

As previously noted, the IIS should offer a variety of methods to support a range of EHR capabilities and preferences. With onboarding, the legacy data request should be discussed early in the onboarding process to improve the chances of receiving data through a legacy load at or near the “go live” date. IIS should also consider a more lenient data quality approach to legacy submissions to avoid placing unrealistic expectations on providers and their EHR vendors to correct legacy records.

¹⁷ The ability to extract and transmit legacy data from EHRs may become more commonplace if the rule proposed by DHHS to improve the interoperability of health information is implemented. IIS programs should monitor this rule as discussions and activities regarding this issue further evolve. <https://www.healthit.gov/topic/laws-regulation-and-policy/notice-proposed-rulemaking-improve-interoperability-health>.

¹⁸ The ability to readily export all patients from one EHR to another when a provider changes its health IT system is a core focus of the NPRM to improve the interoperability of health information. If enacted, access to a provider's data from a previous EHR platform should no longer be a barrier to sharing legacy data with an IIS.

Most EHRs should be able to identify and transmit legacy data using one or more of the methods proposed in the earlier chapter on [legacy import methodologies](#). Any costs imposed by EHR vendors to perform these activities are likely a reflection of the man-hours to help identify the records of interest and facilitate the transmission to the IIS. If custom development work is proposed, the IIS should work with the EHR vendor to identify other strategies that may require less effort to achieve the same outcome.

In situations where the provider has transitioned to a new EHR platform and did not migrate its existing data, there may be little that the IIS or EHR vendor can do to procure the provider's data without expending significant resources. Some IIS programs have suggested offering a reverse data load from the IIS to the provider's new EHR to help seed the patient and vaccination data based on records previously reported by the provider to the IIS as a gesture of goodwill. Whenever the IIS program learns that a provider is transitioning to a new EHR, the IIS should encourage the provider to request a migration of its existing data to its new platform or attempt to secure a legacy vaccination load before the termination of support on its existing system.



Implementation Considerations and Recommendations: Priorities and Barriers

- Assess the current process for pursuing legacy data and remove any unnecessary requirements/barriers.
- Promote the importance of securing legacy data and demonstrate the value of these activities to the provider community.
- When a provider is changing EHR products, encourage the provider to migrate patient and vaccination data to the new system or try to secure an immunization data extract from the old system if possible.

2.6 ALTERNATIVE DATA SOURCES

IIS programs may also want to explore alternative sources of vaccination data to improve overall IIS data saturation. While a provider or vaccination clinic is always the most reliable primary data source, other secondary data partners may have valuable information to fill in IIS data gaps or to help populate less mature IIS databases. Alternative data partners may also be a good source of adult vaccination data. Suggested data partners include:

- Medicaid
- Medicare
- Insurance companies/large health plans
- Pharmacies
- Very large organizations, like Kaiser Permanente and Ochsner
- Regional public health and community health centers.

Of these options, **Medicaid is considered to be the most viable and significant source of administered vaccination data**, especially in states with large Medicaid populations. In the past, Medicaid data were considered of lower

quality than IIS data; however, this perception has been changing as both CMS and IIS strategies for addressing data quality have improved over the years. Medicaid data can provide additional benefits to an IIS, such as Medicaid number, updated demographic/contact information (mailing address and email address), and clinical/medical home. Other insurance providers can also be good sources of data but do introduce more data quality issues, especially with date of administration versus billing date, single antigens versus combos, and use of CPT codes versus CVX codes.

In addition, many IIS programs would like to do more data sharing with schools but are significantly restricted by FERPA¹⁹ consent rules. Some IIS programs offer to help populate school system databases, but schools are unable to share data back to the IIS without student or parent/guardian consent.²⁰ School-based health clinics are gaining popularity and are creating new opportunities for the sharing of student vaccination records.

¹⁹ Family Educational Rights and Privacy Act: <https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html>.

²⁰ Data quality in school-based immunization records may be a concern for some jurisdictions based on reporting source and data collection practices. IIS programs should employ all standard onboarding processes and protocols when establishing an electronic interface with any student management system. These records should always be marked as historical unless administered through a school-based health clinic.

Implementation Considerations and Recommendations: Alternative Data Sources

- Explore opportunities to partner with other significant data sources to improve data saturation or fill in routine data gaps.

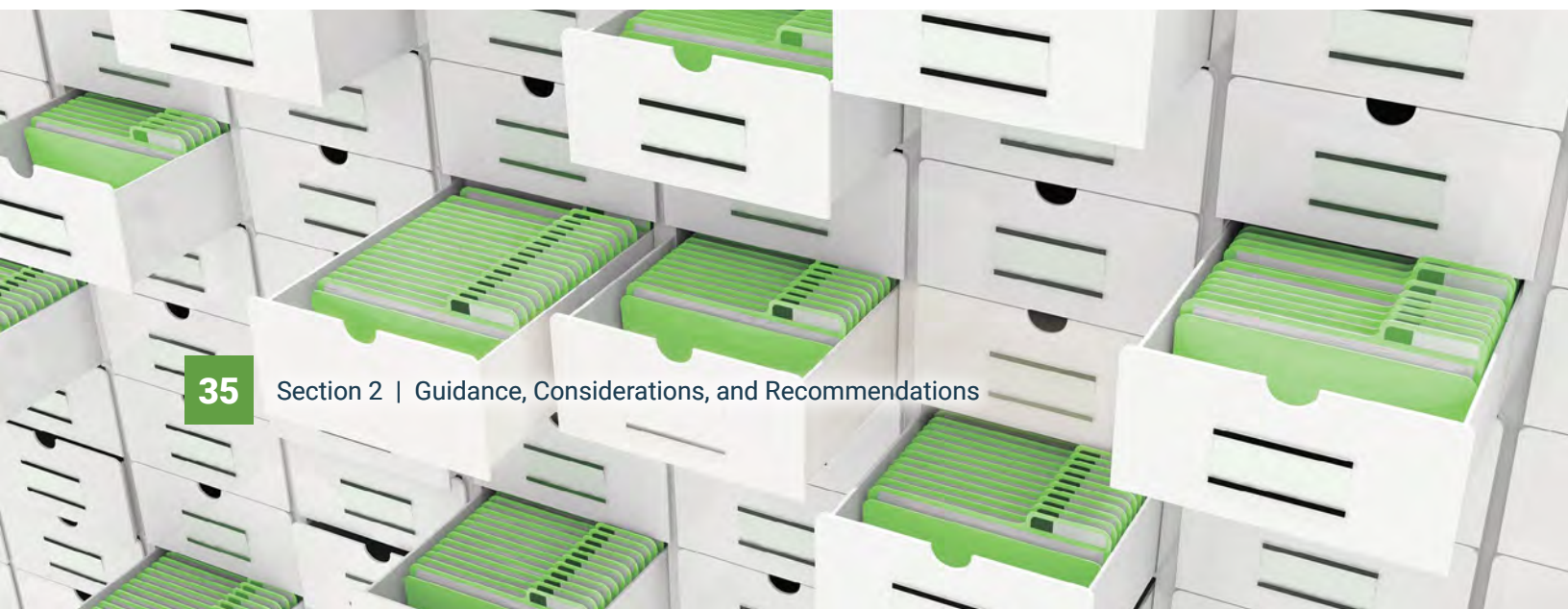
2.7 GAP MONITORING

Missing data in the IIS production environment can result from gaps in reporting or incomplete patient histories. **It may be difficult for an IIS program to determine which patients and vaccinations are missing from the IIS because they were not reported versus doses that were simply never administered.** Most often, gaps are identified accidentally as IIS staff or providers happen upon a missing record or vaccination event. Some IIS programs have established routine processes for gap monitoring by leveraging IIS reports, back-end queries, coverage assessments, and HL7 interface monitoring tools.

Most IIS platforms offer **reports** that can be used to identify patients missing one or more vaccinations, such as a missing immunizations report or reminder/recall patient listing. In order to confirm whether missing doses were administered, providers should be asked to compare these lists against patient records in the EHR or paper charts. This can be a time-intensive activity, so IIS programs may have more success if this is recommended to providers as a monthly maintenance activity.

IIS programs can also leverage **back-end queries** to identify patients with incomplete records and general trends in provider vaccination and reporting habits. These in-depth queries can be used on a routine basis or as needed to diagnose broader behaviors. This information can then be used to develop training and intervention strategies.

IIS-based coverage assessments, like IQIP and routine assessment activities, can be another indicator of missing vaccination data. If a provider's coverage rate seems artificially low, the provider can review missing immunization reports (numerator) and active patient listings (denominator). Quality improvement activities can be used to assess why vaccination information was missing from the IIS, and patient listings can help verify/confirm patients that no longer reside within the jurisdiction.



With providers that report through an HL7 interface, **active monitoring of HL7 traffic** is the easiest method to identify reporting problems. On a daily basis, providers should be reviewing acknowledgement (ACK) messages to confirm that (1) messages are being received by the IIS and (2) there have been no rejections (AR) or errors (AE) with a severity (ERR) of “error” (E). On a monthly basis, providers should confirm that the number of patients reported and doses administered in the practice equals the number of patients and doses received by the IIS.

IIS programs are encouraged to leverage IIS reports, back-end queries, coverage assessments, and HL7 interface monitoring tools to develop routine strategies for identifying gaps in IIS records. Some IIS programs may push the responsibility for gap monitoring down to the provider level. Providers should be trained on the various tools available in the IIS and how to identify and resolve missing data. In any case, **gap monitoring for patients and vaccinations is an opportunity for IIS programs and providers to work together to ensure that the IIS always includes the most complete and up-to-date vaccination records for use by all of the stakeholders that rely on this information.**

Implementation Considerations and Recommendations: Gap Monitoring

- Develop a monitoring plan/process to routinely assess the production IIS and identify gaps in patient and vaccination reporting.

2.8 GAP RESOLUTION

Once gaps have been identified, the IIS should have a protocol for assessing why the issue occurred and a method for acquiring any missing patient and vaccination records. Some issues may be isolated to a single patient, while others may relate to a specific data element (e.g., CVX code or historical/administered flag) or reporting time frame.

Some of the more common reasons that gaps occur include:

- Failure to actively monitor the HL7 interface performance and acknowledgements
- System outages
- Staff turnover
- Code changes in the EHR or IIS
- Delayed data entry into the EHR or reporting of administered doses to the IIS

When the issue relates to a single patient, the provider should correct the record in the IIS or correct the record in the EHR and trigger the message for resubmission to the IIS. For issues that affect a larger number of records, the provider, EHR vendor and/or IIS program should determine the common link between the records (data element or time frame) and resolve appropriately.

- If the issue is related to a **specific data element**, records should be corrected (manually or through the use of a script) and flagged for resubmission.
- If the issue is related to a **specific time frame**, the provider/EHR vendor should establish a time/date parameter to select the affected records and resubmit them to the IIS.
- If the issue relates to **the interface itself**, the interface should be fixed and retested before additional records are submitted to the IIS. Affected records should then be corrected (if applicable) and flagged for resubmission.

Providers should be aware of how to correct data issues in the IIS or their EHR but may need to work with their EHR vendor for resubmitting messages to the IIS. **Most EHR vendors should easily be able to flag and resubmit affected records (using a queued batch protocol).** For smaller product vendors or vendors without these capabilities, it may be easier to resubmit as a full legacy load using one of the methods previously described. If the gaps are the result of a provider's transition to a new EHR platform, it may be both cost and resource prohibitive for the provider to provide missing patient and vaccination data from its legacy system.

Implementation Considerations and Recommendations: Gap Resolution

- Develop a resolution plan/process to address gaps identified in patient and vaccination data or reporting.
- Ensure that providers are aware of how to monitor their data, correct issues, and resubmit messages to the IIS as needed.

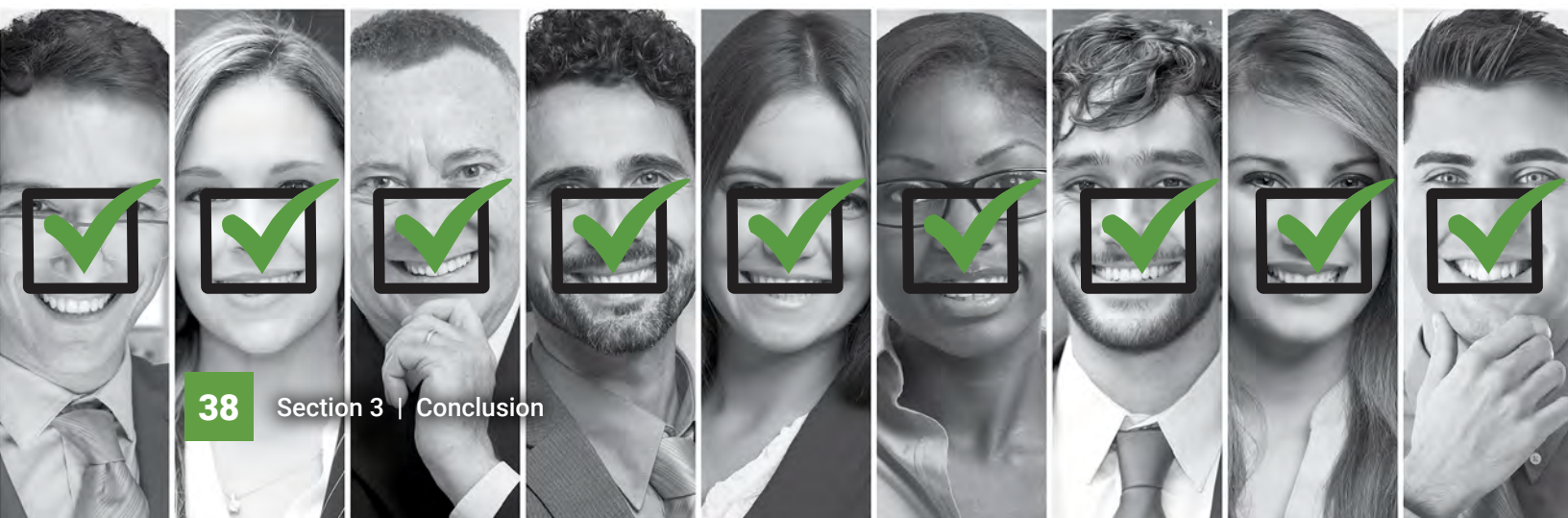
SECTION 3 CONCLUSION

An IIS offers its greatest value when fully populated with patient and vaccination data that can be leveraged by all stakeholders that rely on this information.

Improving data saturation and partnering with providers to optimize record sharing benefits the entire community and promotes the core public health mission of preventing and containing vaccine-preventable disease. Legacy data loads and active monitoring of the IIS have a direct impact on improving IIS data saturation by increasing the amount of complete patient and vaccination data available in the IIS. More complete vaccination records improve vaccine forecasting, prevent duplicate doses, and decrease missed opportunities to vaccinate.

IIS programs should strive to prioritize the pursuit of legacy data and incorporate this activity into the standard onboarding process for all new interfaces. IIS program staff should evaluate their current processes for legacy data collection to identify opportunities for improving and expanding upon these activities. IIS can mediate provider/EHR vendor barriers and challenges by offering a variety of options for transmitting legacy data, modifying data quality expectations for legacy data imports, and promoting the benefits and value of achieving complete records in the IIS.

A consolidated summary of all the implementation considerations and recommendations from each topic section has been included in [Appendix B. Summary of Implementation Considerations and Recommendations](#) for quick reference.





APPENDICES



APPENDIX A ABBREVIATIONS/ACRONYMS	40
---------------------------------------------	-----------

APPENDIX B SUMMARY OF IMPLEMENTATION CONSIDERATIONS AND RECOMMENDATIONS	41
-----------------------------------------------------------------------------------------	-----------

APPENDIX C REFERENCES	43
---------------------------------	-----------

APPENDIX D ACKNOWLEDGMENTS	44
--------------------------------------	-----------

APPENDIX A ABBREVIATIONS/ACRONYMS

Table 3 | *Abbreviations/Acronyms*

ABBREVIATIONS/ACRONYMS	
ACK	HL7 Acknowledgment Message
AIRA	American Immunization Registry Association
CDC	Centers for Disease Control and Prevention
CDS	Clinical Decision Support
CDSi	CDC Clinical Decision Support Logic for immunizations
CVX	CDC Code for Vaccine Administered
DHHS	US Department of Health and Human Services
DQA	Data Quality Assurance
EH	Electronic Health Information
EHR	Electronic Health Record
ERR	An HL7 message segment
FERPA	Family Educational Rights and Privacy Act
HL7	Health Level Seven International
IIS	Immunization Information System
IQIP	Immunization Quality Improvement for Providers (previously AFIX)
IT	Information Technology
MIROW	Modeling of Immunization Registry Operations Workgroup
NDC	National Drug Code
NIS	National Immunization Survey
NOS	Not Otherwise Specified
NPRM	Notice of Proposed Rulemaking
ONC	Office of the National Coordinator for Health Information Technology
SFTP	Secure File Transfer Protocol
SME	Subject Matter Expert
VXU	HL7 Unsolicited Vaccination Record Update Message

APPENDIX B SUMMARY OF IMPLEMENTATION CONSIDERATIONS AND RECOMMENDATIONS

METHODOLOGIES

- Offer a variety of methods to support a range of EHR capabilities and preferences.
- Leverage the IIS-EHR HL7 interface for legacy loads unless there are other prevailing circumstances.

TIMING

- Discuss the legacy data request with each provider early in the onboarding process, preferably during the initial kick-off call.
- Prioritize legacy data loads for the period immediately after onboarding.
- Develop a strategy to secure legacy data from already existing provider interfaces.

VALIDATION

- Apply the same data validation methods and criteria used for routine data submission (e.g., HL7 technical validations).
- Accept as much viable data as possible even if legacy records do not conform to current data quality standards.
- Schedule the timing of the legacy load with the provider/EHR vendor so IIS program staff can account for an increased spike in errors and warnings.

DUPLICATES

- A good patient-level and vaccine-level deduplication algorithm for identifying and resolving duplicates is critical for the proper processing of legacy data.

HISTORICALLY “ADMINISTERED”

- Use standard HL7 technical validations for accepting/rejecting records, but if legacy messages flagged as “administered” are likely to fail, consider using historical flags or an alternative method to load (e.g., traditional flat file).

CVX CODES

- Assess CVX code validity based on date of vaccine administration by maintaining the CVX code table for all active and inactive CVX codes.

GENERIC DATES AND DATA SOURCES

- Accept records with generic dates or questionable data sources at face value as reported by the provider.

PERFORMANCE

- Coordinate the timing of the legacy load with the provider/EHR vendor.
- Break up the transmission by a selected parameter, record count, or file size if needed.

PRIORITIES AND BARRIERS

- Assess the current process for pursuing legacy data and remove any unnecessary requirements/barriers.
- Promote the importance of securing legacy data and demonstrate the value of these activities to the provider community.
- When a provider is changing EHR products, encourage the provider to migrate patient and vaccination data to the new system or try to secure an immunization data extract from the old system if possible.

ALTERNATIVE DATA SOURCES

- Explore opportunities to partner with other significant data sources to improve data saturation or fill in routine data gaps.

GAP MONITORING

- Develop a monitoring plan/process to routinely assess the production IIS and identify gaps in patient and vaccination reporting.

GAP RESOLUTION

- Develop a resolution plan/process to address gaps identified in patient and vaccination data or reporting.
- Ensure that providers are aware of how to monitor their data, correct issues, and resubmit messages to the IIS as needed.

APPENDIX C REFERENCES

1. AIRA. Consolidating Demographic Records and Vaccination Event Records. Modeling of Immunization Registry Operations Workgroup (MIROW). 2017. <https://repository.immregistries.org/resource/consolidating-demographic-records-and-vaccination-event-records/>
2. AIRA. Data Quality Assurance: Selected Aspects. Modeling of Immunization Registry Operations Workgroup (MIROW). May 2013. <http://repository.immregistries.org/resource/data-quality-assurance-in-immunization-information-systems-selected-aspects/>
3. AIRA. Data Quality Assurance in Immunization Information Systems: Incoming Data. Modeling of Immunization Registry Operations Workgroup (MIROW). February 2008. https://repository.immregistries.org/files/resources/5835adc2dbbe4/data_quality_assurance_in_immunization_information_systems_incoming_data_.pdf
4. AIRA. Data Validation Guide for the IIS Onboarding Process. 2017. http://repository.immregistries.org/files/resources/58a601d626d7a/aira_data_validation_guide_-_final_new_logo.pdf
5. AIRA. IIS Data Quality Practices: Monitoring and Evaluating Data Submissions. September 2017. http://repository.immregistries.org/files/resources/59cabe6404421/data_quality_phase_ii.pdf
6. AIRA. IIS Data Quality Practices: To Monitor and Evaluate Data at Rest. May 2018. https://repository.immregistries.org/files/resources/5c002cbde216d/aira_dq_guide_data_at_rest_-_final.pdf
7. AIRA. Onboarding Consensus-Based Recommendations. November 2018. https://repository.immregistries.org/files/resources/5c377a4b2a490/aira_onboarding_recommendations_final.pdf
8. AIRA. Vaccination Level Deduplication in IIS. Modeling of Immunization Registry Operations Workgroup (MIROW). December 2006. <https://repository.immregistries.org/resource/vaccination-level-deduplication-in-immunization-information-systems-1/>
9. CDC. CDC-Endorsed Data Elements, Version 4. Accessed September 23, 2018. <https://www.cdc.gov/vaccines/programs/iis/core-data-elements.html>
10. CDC. Clinical Decision Support for Immunization (CDSi). Accessed April 15, 2019. <https://www.cdc.gov/vaccines/programs/iis/cdsi.html>
11. CDC. CPT Codes Mapped to CVX Codes. Accessed April 15, 2019. <https://www2a.cdc.gov/vaccines/IIS/IISStandards/vaccines.asp?rpt=cpt>
12. CDC. HL7 Version 2.5.1 Implementation Guide for Immunization Messaging, v 1.5, November 2014, and Addendum. July 2015. <https://www.cdc.gov/vaccines/programs/iis/technical-guidance/hl7.html>
13. CDC. IIS: Current HL7 Standard Code Set – CVX – Vaccines Administered. Accessed April 15, 2019. <https://www2a.cdc.gov/vaccines/IIS/IISStandards/vaccines.asp?rpt=cvx>
14. CDC. IIS Functional Standards, v4.0. Accessed September 23, 2018. <https://www.cdc.gov/vaccines/programs/iis/func-stds.html>
15. CDC. Immunization Information Systems Patient-Level De-Duplication Best Practices. June 25, 2013. <https://www.cdc.gov/vaccines/programs/iis/interop-proj/downloads/de-duplication.pdf>
16. CDC. Logic Specification for ACIP Recommendations v3.0. April 2016. <https://www.cdc.gov/vaccines/programs/iis/interop-proj/downloads/logic-spec-acip-rec.pdf>

APPENDIX D ACKNOWLEDGMENTS

AIRA would like to acknowledge and thank the following individuals and organizations for their support and assistance:

- Individual subject matter experts who contributed their expertise through telephone interviews and document review:
 - **Michael Powell**, California
 - **Kim Gulliver**, Colorado
 - **Steve Jarvis**, Colorado
 - **Jody Schweitzer**, Kentucky
 - **Laura Barrett**, Kentucky
 - **Rashid Malik**, Maryland
 - **Patricia Swartz**, Maryland
 - **Kurt Seetoo**, Maryland
 - **Alok Patra**, New Jersey
 - **Monica Sull**, New York City
 - **Alexandra Ternier**, New York City
 - **Vikki Papadouka**, New York City
 - **Megan Meldrum**, New York State
 - **David Baron**, Tennessee
 - **Melissa Davis**, Tennessee
 - **Kevin Allen**, Texas
 - **David Krenek**, Texas
 - **Matthew Davis**, Texas
 - **Danny Wise**, Allscripts
 - **Michelle Sumner**, Allscripts
 - **Pat Quinlan**, Allscripts
 - **Greg Faber**, Epic
 - **Emily Smith**, Epic



- The AIRA board of directors who provided input at various stages of the effort and/or reviewed and provided comment on the final guide:
 - *President* • **Amanda Harris**, Nevada Department of Health
 - *President-Elect* • **Aaron Bieringer**, Minnesota Immunization Information Connection
 - *Immediate Past President* • **Kim Salisbury-Keith**, Rhode Island Department of Health
 - *Secretary* • **Jenne McKibben**, Oregon Immunization Program
 - *Treasurer* • **Heather Roth**, Colorado Department of Public Health and Environment
 - Directors:
 - **Bridget Ahrens**, Vermont Immunization Registry
 - **Christy Gray**, Virginia Department of Health
 - **Dannette Dronenberg**, Washington State Department of Health
 - **David McCormick**, Indiana State Department of Health
 - **Jeffrey McIntyre**, Mississippi State Department of Health
 - **Kevin Dombkowski**, University of Michigan, Child Health Evaluation and Research Unit
 - **Nathalie Hartert**, Tennessee Department of Health
 - **Steve Murchie**, Envision Technology Partners
- The AIRA staff who contributed to this document's development:
 - **Mary Beth Kurillo, MPH, MSW**, AAIRA Policy and Planning Director, who provided leadership oversight, provided guidance and input, and edited the guide
 - **Nichole Lambrecht**, AIRA Senior Project Manager, who oversaw and managed the project, provided guidance and input, and edited the guide
 - **Danielle Reader-Jolley**, Independent Public Health Consultant, who researched, wrote, and edited the guide
- An independent copy editor, **Maureen Brody**, who reviewed and professionally edited the guide.
- Individuals, in addition to the subject matter experts, who provided feedback during the review process:
 - **Craig Newman**, Altarum
 - **Celena Calvo-Story**, Guam
 - **Noam Arzt**, HLN Consulting
 - **Christopher Jay**, Kansas
 - **Brandy Altstadter**, Scientific Technologies Corporation
 - **Bridget Ahrens**, Vermont
 - **Carolina Danovaro**, World Health Organization