PathData Use and Sharing Agreement to Support the United States Government's COVID-19 Emergency Response
Jurisdiction Immunization and Vaccine Administration Data Agreement

This Data Use and Sharing Agreement ("DUA") is made between Missouri Department of Health and Senior Services, located at 930 Wildwood Drive, Jefferson City, MO 65109 and the Centers for Disease Control and Prevention (CDC), an agency of the Department of Health and Human Services ("HHS"), to describe the data use and sharing parameters for certain immunization and vaccine administration data, as further described herein. This DUA: 1) describes platforms for the rapid collection, transmission, use, storage, and maintenance of these data available to data sources, CDC, and other users; 2) establishes the terms and conditions for the sharing, protection, and use of these with CDC, HHS and other federal partners; and 3) sets forth the roles and responsibilities of each party.

The DUA is effective as of October 23, 2020.

Background and Purpose

Access to immunization and vaccine administration data is critical to the whole of government response to the Coronavirus Disease 2019 (COVID-19) public health emergency. In furtherance of federal government response efforts, HHS and CDC seek to obtain and utilize these data from various immunization and vaccine data sources, including a jurisdiction's immunization information system (IIS), pharmacies, federal Provider Organizations, and other relevant parties for a range of purposes, including but not limited to rapidly assessing patterns of vaccination among the population; identifying pockets of undervaccination ; assisting in determining vaccine resource allocation to address the needs of jurisdictions; monitoring vaccine effectiveness and safety; assessing spectrum of illness, disease burden, risk factors for severe disease and outcomes; and helping to understand. the impact of COVID-19 on the healthcare system and communities.

To support these purposes, HHS and CDC: 1) have made available a platform for use by data sources to manage, share, and store their immunization data; 2) have developed platforms for use by HHS, CDC, and other federal partners to extract, accept, manage, share, and store relevant immunization data in furtherance of the response; 3) will, consistent with applicable law, enable the secure transmission of extracted data from and across these platforms for further use by a jurisdiction, CDC, HHS, and other federal partners in furtherance of the response; 4) as applicable, will assure compliance of these platforms with the Federal Information Security Management Act (FISMA) and other federal data security policies; and 5) will provide operational support to the data sources and other authorized users of the various platforms, as appropriate.

Authority

HHS and CDC are authorized by Sections 301 and 319D of the Public Health Service Act [42 U.S.C. §§ 241 and 247d-4], as amended, to maintain active surveillance of diseases through epidemiologic and laboratory investigations and data collection, analysis, and distribution. The Jurisdiction entering this DUA agrees that it is authorized to send the Covered Data to and through the COVID-19 Clearinghouse, the Immunization (IZ) Data Lake, the Vaccine Administration Management System (VAMS), and HHS Protect/Tiberius, as those platforms and systems are further defined herein, and/or will obtain consent from any external entities or individuals from whom it collects data to allow for such sharing and use.

In addition, HHS and CDC each is a "public health authority" as defined at 45 C.F.R. §164.501 and as used in 45 C.F.R. §164.512(b), Standards for Privacy of Individually Identifiable Health Information, promulgated under the Health Insurance Portability and Accountability Act of 1996 ("HIPAA") and, as such, are authorized by 45 CFR 164.512(b) to receive Protected Health Information ("PHI").

As applicable, the Parties acknowledge that Jurisdiction may be a hybrid entity for purposes of HIPAA. Jurisdiction's healthcare component is neither involved nor implicated in this DUA. For purposes of this DUA, Jurisdiction is a public health authority under 45 CFR §164.512 and is neither a covered entity nor a business associate, as defined under 45 CFR §160.103. The Parties expressly do not intend to create a HIPAA business associate relationship, and nothing in this DUA may be construed to make Jurisdiction a covered entity or business associate for purposes of this agreement.

Data Use and Sharing Terms

1. Platforms

In furtherance of the activities set out in this DUA, HHS and CDC, either directly or by and through a service provider, have stood up, are supporting, and/or are expanding the capacity of the following secure, certified, cloud-based data management platforms:

a. Immunization (IZ) Gateway: The IZ Gateway is a cloud-hosted message routing service offered by the Association for Public Health Laboratories (APHL) and is intended to enable data exchange across IISs, other provider systems, the COVID-19 Clearinghouse, and the IZ Data Lake. A jurisdiction may enter into appropriate agreements with APHL to enable its IIS to update, query, and report immunization data to and through the IZ Gateway. The IZ Gateway is intended to allow a jurisdiction to connect its IIS and other provider systems to the COVID-19 Clearinghouse and the IZ Data Lake; connect its IIS to VAMS data, where applicable; enable queries and route messages to/from its IIS; and route secured, standardized HL7 messages from its IIS, VAMS, or other provider organizations to the COVID-19 Clearinghouse and the IZ Data Lake.

   A jurisdiction choosing to use the IZ Gateway will enter into relevant agreements with APHL with respect to use of the IZ Gateway.

b. COVID-19 Clearinghouse: The COVID-19 Clearinghouse is a cloud-hosted data repository provided and managed by HHS that, as a functional tool, provides a secure space for a jurisdiction to upload and store COVID-19 vaccination data collected from provider organizations via electronic health records (EHRs) and from pharmacy systems. The COVID-19 Clearinghouse is intended to allow a jurisdiction to upload, store, reconcile, and manage data for general COVID-19 vaccine administration, to meet reporting requirements and needs, and to allow providers to search for a patient, see what brand of COVID-19 vaccine they received, and when they received their first dose of COVID-19 vaccine to ensure dose matching to complete the vaccine series (see Appendix A).

   The COVID-19 Clearinghouse will be able to receive data from VAMS and other vaccination sources not already onboarded as providers to an immunization registry and will allow the data to be shared across relevant jurisdictions for reconciliation of patient registry records. Neither CDC nor HHS, either directly or by and through the cloud service provider, will have access to personally identifiable vaccination data submitted by a jurisdiction's IIS, VAMS, or other vaccine source record system during transport or record processing or while in storage without the express consent of the jurisdiction. HHS and CDC, either directly or by and through the service provider, will take all reasonable measures to secure the data residing in the COVID-19 Clearinghouse and, except as may be required by applicable federal law, will not, without the prior written authorization from the jurisdiction, further use, disclose, or transmit the data beyond what is described within this DUA. Should disclosure or transmission of a jurisdiction's data be required by federal law, HHS and/or CDC will promptly notify the jurisdiction in writing of the required release.

c. Immunization (IZ) Data Lake: The IZ Data Lake is a CDC secure, cloud-hosted data repository created to receive and store redacted COVID-19 vaccination data for doses administered, coverage, inventory, and distribution. The IZ Data Lake will receive the data from various data flows, including internal CDC sources (VTrckS), Provider Agreement data via an upload portal, VaccineFinder, and the COVID-19 Clearinghouse. These data will be used by CDC to provide aggregate-level reports for COVID-19 vaccine administration, ordering, inventory, and provider information. The IZ Data Lake will also aggregate and analyze data and provide data summaries and analytics via platforms such as the Data Storefront HHS Protect, and HHS Tiberius.

d. VaccineFinder: The VaccineFinder website (www.vaccinefinder.org) helps people find providers who offer specific vaccines. VaccineFinder will serve two roles during the COVID-19 Vaccination Program:

1. **Inventory reporting**: Approved COVID-19 vaccination providers will report on-hand COVID-19 vaccine inventory daily.
2. **Increase access to COVID-19 vaccines**: COVID-19 vaccination providers may choose to make their location(s) visible on VaccineFinder to increase access to COVID-19 vaccines once supply is available for the general population.

VaccineFinder will exchange data with the IZ Data Lake for the purposes of provider pre-enrollment, data analysis, and summaries via platforms such as the Data Storefront, HHS Protect, and HHS Tiberius.

e. VTrckS: CDC's Vaccine Tracking System is the platform for ordering all COVID-19 vaccines. VTrckS users will use the system to:

- View vaccine allocations allotted to each program.
- Place and/or manage vaccine orders for their providers.
- Generate reports throughout the vaccine distribution process, from placing vaccine orders through distribution.

VTrckS receives data from jurisdiction immunization registries and transmits data to these registries, VaccineFinder, and the IZ Data Lake.

f. Provider system: A provider system is any platform used by a vaccination provider to track the administration and uptake of vaccine among their patient populations. Vaccination providers generally utilize an Electronic Health Record (EHR) to connect directly to IISs for vaccine uptake tracking. In some instances, vaccination providers will utilize a direct user interface (UI) to enter vaccination information directly into an IIS UI portal.

g. Other vaccine source: These are alternative sources (in addition to IISs and EHRs) that may store vaccination information about an individual. Examples include travel vaccination records, passports, vital records/birth records, Medicaid records, insurance claim information, etc. and their corresponding systems.

h. HHS Tiberius: Tiberius provides a COVID-19 vaccine distribution planning, tracking, modeling, and analysis ecosystem. Tiberius leverages the same technologies as the HHS Protect Platform ("HHS Protect") and integrates data sources from federal agencies, state and local partners, private sector partners, and open data providers to create a comprehensive common operating picture for the COVID-19 vaccine planning, distribution, and administration effort.

2. Definitions:
For purposes of this Agreement, the following definitions shall apply and may be used in the main body of the DUA and/or in relevant appendices:

"Authorized User," for purposes of this DUA, means an individual who, as part of directly supporting the whole of government response efforts, has a need for data stored in the COVID-19 Clearinghouse, the IZ Data Lake and/or the Tiberius platforms in furtherance of the purposes and uses set forth herein. Authorized Users will generally be employees, contractors, and/or other agents specified by Jurisdiction or federal agencies engaged in the response for purposes of addressing critical public health and emergency response activities, including assessing infrastructure needs and resource allocation. Authorized Users must adhere to applicable federal law and, as consistent and applicable the provisions set out in this DUA with respect to the data stored in the respective platforms.

"Data Source," for purposes of this DUA, is a Jurisdiction which, by and through an IIS and/or similar system(s) created to serve a range of administrative functions related to vaccines, provides Covered Data as set forth herein. Generally, the IIS or related system(s) will collect

data from public and private health care provider organization e.g. EHRs, health information systems, (e.g., vital statistics, state Medicaid agencies, etc.), and pharmacies.

"Covered Data" means the information that is being shared by the Data Source with each relevant platform as further described in Appendices A-D, but that is generally categorized into four _primary datasets: the VAMS data, the COVID-19 Clearinghouse Data, the IZ Data Lake Data, and the Tiberius Data. HHS and CDC acknowledge that the Covered Data to which each agency will have access is the minimum amount of information necessary to accomplish public health or emergency response needs. A list of Covered Data elements for each dataset is provided in Appendices A-D.

Covered Data may be used by Authorized Users within the parameters set forth in this DUA. The data elements listed in Appendices A-D will be updated periodically as more information on COVID-19 immunization is available. The overall DUA will remain unaffected by subsequent updates. Appendices A-D also provide the mode and method of secure transmission of the data from the Jurisdiction's IIS or similar system(s) directly to the COVID-19 Clearinghouse; from the COVID-19 Clearinghouse to the IZ Data Lake; and from the IZ Data Lake to Tiberius. This information includes the potential availability and use of a privacy-preserving record linkage (PPRL) tool, which may be made available by HHS or CDC, either directly or by and through a contractor (Appendix E). Of note, data entering the COVID-19 Clearinghouse through the IZ Gateway will be governed by agreements between the IIS jurisdiction and APHL.

"Jurisdiction" means the state, territorial or local health jurisdiction operating under either statutory or regulatory authority to obtain and use health-related data for population health protection. For the purposes of this document, Jurisdictions are funded under CDC-RFA-IP19-1901 317 Notice of Funding Opportunity.

"Immunization Information System" or "IIS are confidential, population-based, computerized databases that record all immunization doses administered by participating providers to persons residing within a given geopolitical area.

"Deidentified Data" means data that do not identify an individual and there is no reasonable basis to believe the information can be used to identify an individual because the data have been rendered not identifiable in accordance with the HIPAA standards set forth in 45 CFR §164.514.

"Party" means a state, territorial or local jurisdiction or CDC; "Parties" means state, territorial, or local jurisdiction and CDC.

"Privacy-Preserving Record Linkage (PPRL)" means the process whereby personally identifiable information (PII) is redacted from a patient/customer record using a one-way, irreversible encryption algorithm to create one or more unique tokens that replace PII elements and allow data systems to match patient/customer records. PPRL is an industry standard that has been implemented and integrated across several data collection sectors where an individual's privacy must be maintained (e.g., health care, biomedical research, payment and claims, retail, intelligence, social research, and public health). For COVID-19 immunization reporting, PPRL offers jurisdictions a mechanism to meet applicable jurisdiction regulations where data sharing with partners such as HHS and CDC may be limited.

"Vaccine Administration Management System" or "VAMS" means the CDC-provided and supported web-based application that provides an option for a jurisdiction to plan and execute COVID-19 vaccine administration in a mass vaccination setting. VAMS has four users with multiple roles within each user module: 1) jurisdictions can provide end-to-end mass vaccination capability and manage mass vaccination clinics; 2) healthcare providers can manage patient scheduling, vaccine administration workflow, and patient monitoring, support social distancing requirements with a scheduling feature, track vaccine inventory and usage, and include warnings when inventory is low; 3) employers/organizations can bulk input employees who will receive an email to register in VAMS; and 4) vaccine recipients can schedule vaccination appointments and receive appointment reminders. For purposes of this DUA, VAMS may be used to send data directly to the COVID-19 Clearinghouse or through the IZ Gateway either back to the Jurisdiction IIS and/or to the COVID-19 Clearinghouse. The Jurisdiction IIS may choose to use VAMS or an alternate mechanism (e.g., state-based vaccination clinic solution) to transmit the data to the IZ Gateway and/or the COVID-19 Clearinghouse.

3. Description of Data Requested and Transmission

Data Source agrees to provide data as described in Appendices A-D to and through platforms as indicated therein, subject to the terms and conditions included in this DUA and applicable to that option.

4. Data Use Terms

The Data Source acknowledges and agrees that HHS, CDC, and Authorized Users may use the Covered Data transmitted to the various platforms as described in this DUA and Appendices A-D in furtherance of response activities related to the COVID-19 pandemic. This includes, at a minimum, the following activities:

a. Analyze and visualize the Covered Data to which they have access to improve the monitoring of vaccine and vaccine-related activities related to the COVID-19 pandemic response including vaccine safety and assessment of vaccine effectiveness;
b. Analyze and visualize the Covered Data to improve the monitoring of vaccine safety and assessment of vaccine effectiveness;
c. As applicable to the platform, share the Covered Data and analyses thereof with official federal, state, local, tribal, and territorial government health agencies or other agencies and entities conducting their public health and vaccine response responsibilities consistent with applicable federal law and the terms of this DUA;
d. Develop analytic methods using the Covered Data to identify immediate public health events or concerns at the federal, state, territorial and local level that warrant further public health investigation or immediate public health intervention actions;
e. Enable Authorized Users, including public health and emergency response officials, to query the Covered Data within the HHS and CDC-provided data platforms as may be necessary to carry out critical public health functions;
f. Share specified data elements with HHS Tiberius for the visualization of Vaccine Administration Data; and
g. Publish findings and conclusions related to their analyses of the data provided. As appropriate, publications will acknowledge Data Source as the source of the data in any such

publication. Given the emergent nature of the response, HHS and CDC may not be able to inform or seek approval from Data Source for such publications but will coordinate as soon as possible and practicable.

5. Data Confidentiality and Security

As applicable to the platform, HHS and CDC will establish appropriate administrative, technical, procedural, and physical safeguards to assure the confidentiality and security of Covered Data in their custody and control, consistent with federal requirements under the FISMA) and other applicable federal laws. The safeguards shall provide a level and scope of security that is not less than the level and scope of security established by applicable law for the type of data provided under this DUA. Where Covered Data provided pursuant to this DUA are identifiable or potentially identifiable, CDC agrees to maintain the confidentiality of the Covered Data to the fullest extent required by applicable law, which includes, as applicable, the Privacy Act of 1974; standards promulgated pursuant to), and the Freedom of Information Act (FOIA), including exemptions provided thereunder.

Where required by law and/or where practicable, HHS and CDC agree to notify Data Source before releasing Covered Data to a third party pursuant to a judicial, governmental, or other request under law, to allow Data Source the opportunity to state any objection to the disclosure of the Covered Data.

Transmission of the Covered Data by and through the various platforms in the control of HHS and CDC shall be done in accordance with acceptable practices for ensuring the protection, confidentiality, and integrity of the contents. Covered Data will be maintained and stored in compliance with CDC's security policies and procedures and consistent with applicable law.

Miscellaneous

1. Data Disposition: Data that have been provided to HHS and CDC under this DUA will be archived, stored, protected, or disposed of in accordance with relevant federal records retention requirements.
2. Funding: This DUA is not an obligation or a commitment of funds, or a basis for a transfer of funds, and does not create an obligation or commitment to transfer data, but rather is a statement of understanding between the parties concerning the sharing and use of covered data. Expenditures by each party are subject to its budgetary processes and to the availability of funds and resources pursuant to applicable laws, regulations, and policies.
3. Settlement of Disputes: Disagreements between the parties arising under or relating to this DUA will be resolved by consultation between the parties and referral of the dispute to appropriate management officials of the parties whenever possible.
4. Applicable Laws: U.S. federal law shall govern the construction, interpretation, and performance of this Agreement.

Term of Agreement, Amendment, and Termination:

1. The term of this DUA shall be one year commencing from the date of the final signature or the duration of the national emergency. The DUA may be renewed upon mutual written consent of the parties.
2. Except as otherwise expressly provided herein, this DUA may be amended only by the mutual written consent of the authorized representatives for each party.
3. This DUA may otherwise be terminated with ninety days' advance notice upon written notice by either party.
4. Any notice required under this DUA must be in writing and sent by electronic mail with written acknowledgement of receipt to the email address for each party provided below.
5. Each party represents that the individual signing below on behalf of the party has the authorization to bind the party indicated to this DUA. This DUA may be signed in counterparts and signatures provided electronically will be deemed originals.

CENTERS FOR DISEASE CONTROL AND PREVENTION AND DATA SOURCE

By: Megan C. Lindley
Name: Megan C. Lindley
Title: Acting Associate Director for Science, CDC/NCIRD/ISD
Date: 11/07/2020
Email: MLindley@cdc.gov

By: Marcia Mahaney
Name: Marcia Mahaney
Title Director, Division of Administration
Date: 10/28/2020
Email: Marcia.Mahaney@health.mo.gov

Appendix A: Covered Data–CDC IIS Data Elements for COVID-19 Vaccine Monitoring

A strong, nationally coordinated approach is critical to collecting, tracking, and analyzing vaccination data, especially in early phases of vaccine administration, which is expected to occur in nontraditional settings during this COVID-19 response. This document outlines the anticipated data elements that will be reported to various platforms supported by HHS and CDC, ultimately with certain elements being shared with HHS and CDC. The required data elements in this document represent demographic and vaccination information routinely captured by an IIS during a vaccination event.

Discrete Data Elements

In order to ensure appropriate vaccine administration and distribution by HHS and CDC to jurisdictions and, ultimately, to providers, certain data elements must be provided to HHS and CDC.

Table 1 includes each data element a Jurisdiction will report to VAMS (if applicable), to the IZ Clearinghouse, and to the CDC IZ Data Lake. Table 2 includes each data element that will be optional for submission to each platform. Optional data requirements will support additional national coverage analysis and vaccination monitoring efforts.

Any identifiable data elements will be used to facilitate deduplication of data within the COVID-19 Clearinghouse, an analytic environment that will be used to consolidate, deduplicate, and reconcile vaccine administration information from multiple sources (e.g., jurisdictional immunization programs, pharmacies, Department of Defense, Veterans Affairs, Bureau of Prisons, Indian Health Service). Directly identifiable elements will not be sent or stored in the CDC IZ Data Lake environment. The IZ Data Lake will aggregate and analyze data and will provide data summaries and analytics via platforms such as the Data Storefront, HHS Protect, and the HHS Tiberius systems. Data flowing from the IZ Data Lake to Tiberius will be used by Operation Warp Speed members to track progress of vaccine distribution, provider inventory, and administration.

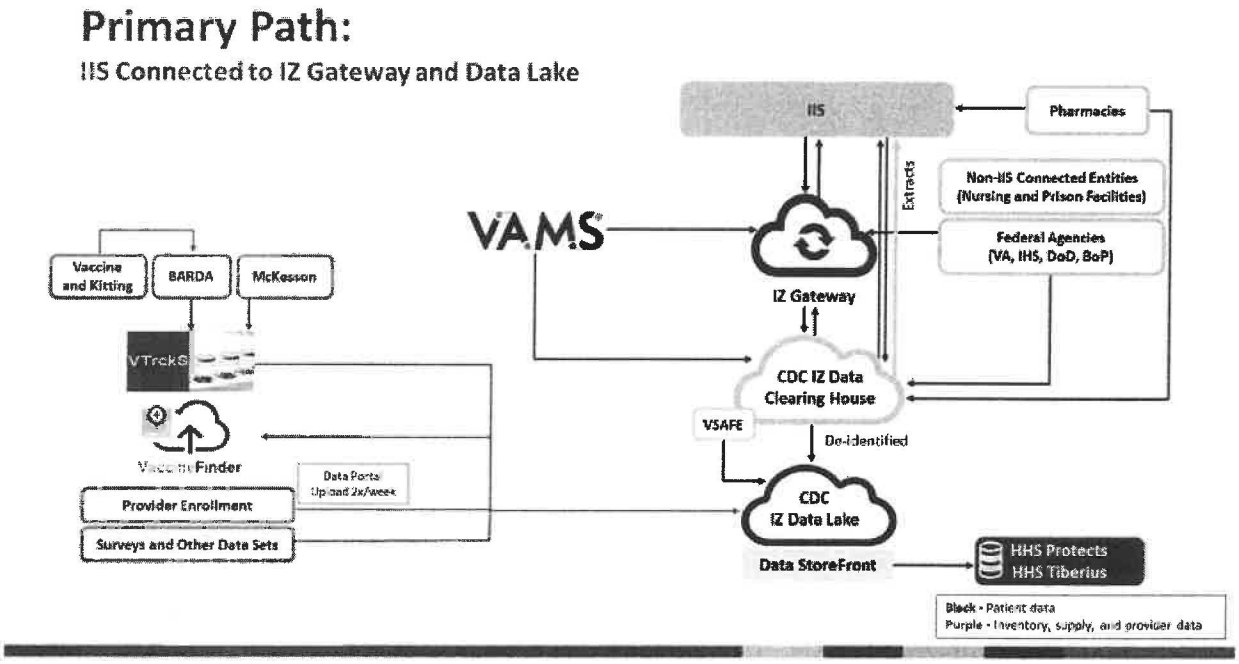
Table 1. Required Data Elements

Required Data Element	VAMS	COVID-19 Clearinghouse	Immunization Data Lake
Data elements required for reporting by IIS			
Administered at location: facility name/ID	✓	✓	✓
Administered at location: type		✓	✓
Administration address (including county)		✓	✓
Administration date		✓	✓

CVX (product)		✓	✓
Dose number		✓	✓
IIS recipient ID*		✓	✓
Recipient race		✓	✓
Recipient ethnicity		✓	✓
IIS vaccination event ID		✓	✓
Lot number: unit of use and/or unit of sale		✓	✓
MVX (manufacturer)		✓	✓
Recipient address*		✓	
Recipient date of birth*		✓	
Recipient name*		✓	
Recipient sex		✓	✓
Sending organization		✓	✓
Vaccine administering provider suffix		✓	✓
Vaccine administering site (on the body)		✓	✓
Vaccine expiration date		✓	✓
Vaccine route of administration		✓	✓
Vaccination series complete		✓	✓
Optional Data Elements			
Optional Data Element	VAMS	COVID-19 Clearinghouse	Immunization Data Lake
Comorbidity status	✓	✓	✓
Recipient missed vaccination Appointment (Y/N)	✓	✓	✓
Serology results (Y/N)	✓	✓	✓
Vaccination refusal (Y/N)	✓	✓	✓

**Identifiable Information*

Appendix B: Vaccine Administration Data Workflow



Description of Data Workflow and Transmission

Data Source will provide vaccine administration data (Appendix A) to the COVID-19 Clearinghouse as described in Appendix D, in part, for purposes of deduplication, record matching, and reconciliation. The COVID-19 Clearinghouse will encrypt and store personally identifiable vaccine administration data, allowing secure, role-based access for authorized data users/systems. It will also support multijurisdictional lookup/queries by vaccination providers for validating vaccine type and manufacturer for administration of a second dose. Once processed in the COVID-19 Clearinghouse, certain data elements (see Appendices A and C) from the COVID-19 Clearinghouse Data will be transmitted to the CDC IZ Data Lake platform for additional analyses by CDC and HHS.

Data may be transmitted to the COVID-19 Clearinghouse through the following applications and technology solutions:

- Option 1: Directly, via a Jurisdiction IIS or Mass Vaccination Clinic Application (e.g., VAMS or a state-based vaccination clinic solution)
- Option 2: Via a Jurisdiction IIS or vaccination clinic application through the IZ Gateway provided by APHL
- Option 3: Submission of record-level vaccination and demographic data to the COVID-19 Clearinghouse through a data extract (see Appendix D)

For Data Sources using the IZ Gateway, data from a vaccination clinic application are submitted to the IZ Gateway through a Simple Object Access Protocol (SOAP) Transport in the form of vaccine record updates (VXU) and acknowledgment (ACK) messages. A similar mechanism is used for the submission from the IZ Gateway into the IZ Data Clearinghouse.

Data from both technical solutions will be submitted directly to the COVID-19 Clearinghouse through functionality existing within the technology and will not require direct action on the part of the Data Source. If the Data Source is unable to submit data through the aforementioned technical solutions, Data Source may transmit Covered Data directly in the form of a data extract to the COVID-19 Clearinghouse.

The COVID-19 Clearinghouse will then deliver certain data to the IZ Data Lake. The Data Source will submit information about COVID-19 vaccinations only. The Data Source agrees to provide all required data elements described in Appendix A unless prohibited by state law. Written documentation of these legal restrictions must be provided.

Data from the IZ Data Lake will transmit into HHS Tiberius. The IZ Data Lake connects to Tiberius by using the IZ Data Lake Azure Synapse component. The Tiberius IP addresses are white-listed, and credentials are exchanged over a secure channel. Tiberius credentials in IZ Data Lake Synapse are only provided access to the data approved by CDC programs. Tiberius pulls the data from IZ Data Lake Synapse when it needs to refresh.

Appendix C: Data Dictionary for CDC Vaccine Administration Requirements for COVID-19 Vaccine Monitoring (attached)

Appendix D: COVID-19 Vaccination Reporting Specifications Document (CRVS) (attached)

APPENDIX E: PRIVACY-PRESERVING RECORD LINKAGE

Currently, there is no consistent way for public health jurisdictions to share vaccine administration data with each other, due to two major constraints. First, these data are considered personally identifiable information (PII) and/or protected health information (PHI) and may be subject to a jurisdiction's laws and regulations that may prohibit or limit the sharing of such information outside the jurisdiction. Second, though most people will complete their dose series in the same jurisdiction, some individuals may cross jurisdictions before completion of the series. There are currently no optimal technical solutions to link vaccine administration data electronically across jurisdictional boundaries.

Privacy preserving record linkage (PPRL) provides a practical way for jurisdictions to exchange information on vaccine administration with federal agencies, while preserving and protecting PII and PHI. By creating tokenized, deduplicated data, PPRL can link an individual's COVID-19 vaccination records to inform booster dose delivery decisions without sharing PII across jurisdiction boundaries. PPRL has the potential to optimize vaccination administration efforts by streamlining processes and resolving orphaned data issues within the IIS by providing record linking without the need to exchange PII.

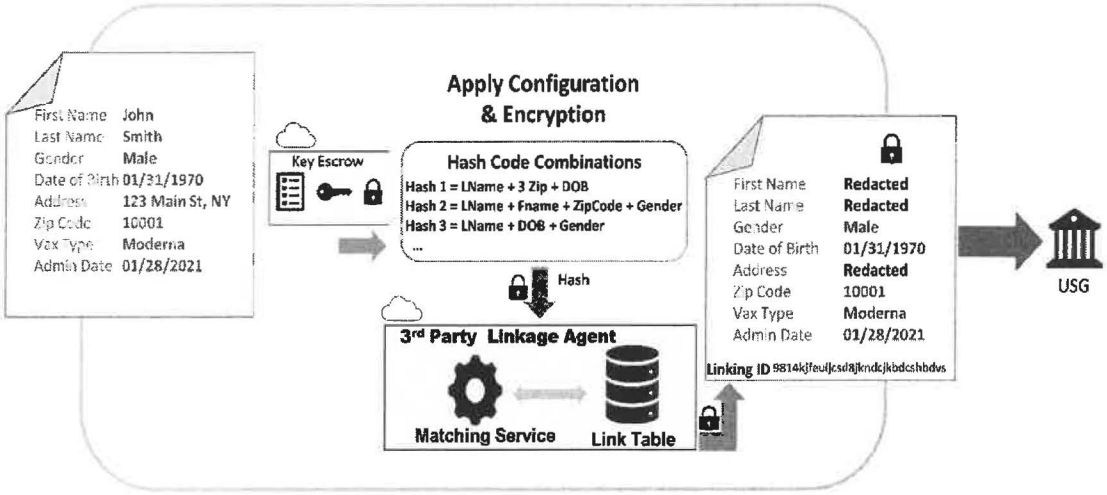
Beyond the dose series look-up use case, PPRL can also enable CDC to associate vaccine administration data from multiple sources (e.g., epidemiologic, laboratory and immunization data) to a specific individual without receiving any PII that might compromise the privacy of that individual.

In the PPRL process, PII is hashed using a one-way, irreversible encryption algorithm to create one or more tokens in a series of prescribed steps prior to transmission beyond an organizational boundary for matching. Hashing works by passing a piece of data through a mathematical function to convert the data into a multi-character code that can be used by computers to match records in the same way that personally identifiable data would be used; the code(s) act as new record identifiers that are used to match with other records similarly converted to codes. The process of hashing results in the creation of unique information based on the PII data of interest that prevents an outside party from recovering the PII, while allowing for the establishment of links across organizations in order to share PII when necessary.

PPRL offers more than one protocol for establishing links. In a "direct" protocol, each party hashes and encrypts their PII and shares the hashed tokens directly with the other party to compare matches. In a "blind" protocol, a third party known as a "linkage agent" is provided access to the hashed data but is unable to view PII. The linkage agent then compares the obfuscated information to establish linkages using the tokens (see Figure below).

Commercial implementations of PPRL services operating in the health domain need to demonstrate that the tokens created are deidentified per the HIPAA standard. This includes demonstrating ability to resist cryptanalytic attacks, dictionary attacks, and statistical attacks (e.g., frequency attacks, collusion attacks).

Figure 1. Tokenization/Unique ID Creation


First Name John
Last Name Smith
Gender Male
Date of Birth 01/31/1970
Address 123 Main St, NY
Zip Code 10001
Vax Type Moderna
Admin Date 01/28/2021
Apply Configuration & Encryption
Key Escrow
Hash Code Combinations
Hash 1 = LName + 3 Zip + DOB
Hash 2 = LName + Fname + ZipCode + Gender
Hash 3 = LName + DOB + Gender
...
Hash
3rd Party Linkage Agent
Matching Service
Link Table
First Name Redacted
Last Name Redacted
Gender Male
Date of Birth 01/31/1970
Address Redacted
Zip Code 10001
Vax Type Moderna
Admin Date 01/28/2021
Linking ID 9814kjfeuljcsdkjkndcjkbdcshbdvs
USG

CDC COVID-19 Vaccine Administration Data Elements

Last Updated 10/5/2020

Required: Data elements required for reporting ***Identifiable information**
Administered at location: facility name/ID
Administered at location: type
Administration address (including county)
Administration date
CVX
Dose number
Extract type
Lot number
MVX
NDC
Recipient address*
Recipient date of birth*
Recipient ethnicity
Recipient ID*
Recipient name*
Recipient race
Recipient sex
Responsible organization
Vaccine administering provider suffix
Vaccine administering site
Vaccine expiration date
Vaccine route of administration
Vaccination series complete
Vaccination event ID

VTrckS provider PIN
Optional: Data elements optional to report
Comorbidity status (Y/N)
Recipient missed vaccination appointment (Y/N)
Serology results (presence of positive result, Y/N)
Vaccination refusal (Y/N)

Data Use and Sharing Agreement: COVID-19 Emergency Response Jurisdiction Immunization and Vaccine Administration Data Agreement Q&A

1. **Can CDC provide more detailed information about breach notification protocols?**

   Transmission of the Covered Data from the jurisdiction to CDC shall be done in accordance with acceptable practices for ensuring the protection, confidentiality, and integrity of the contents. CDC will use all reasonable administrative, technical, and physical measures to safeguard the Covered Data and to protect the Covered Data from unauthorized access, disclosure, use, or modification. This includes setting permissions to access or edit Covered Data commensurate with the level of sensitivity of the Covered Data. Should there be a data breach and unauthorized disclosure of the Covered Data from the CDC controlled platform, a CDC representative shall contact the jurisdiction within 1 hour of the event occurring for personally identifiable information (PII) or protected health information (PHI) and 24 hours for non-PII/non-PHI to make the jurisdiction aware of the situation and provide the response plan.

2. **Can CDC share more detailed information regarding what is meant by an "Authorized User"**
   The current definition of "Authorized User" found on page 3 of the DUA states:

   "For purposes of this DUA, means an individual who, as part of directly supporting the whole of government response efforts, has a need for data stored in the DCH, the IZ Data Lake, and/or the Tiberius platforms in furtherance of the purposes and uses set forth herein. Authorized Users will generally be employees, contractors, and/or other agents specified by jurisdictions or federal agencies engaged in the response for purposes of addressing critical public health and emergency response activities, including assessing infrastructure needs and resource allocation. Authorized Users must adhere to applicable federal law and to any applicable provisions set out in this DUA with respect to the data stored in the respective platforms, which are further defined herein and described in Appendices A–D."

   An Authorized User is an individual, employed by an entity of the government (e.g. CDC, HHS) or their contracted entities (e.g. Oracle for the purposes of the COVID-19 Data Clearinghouse) who is a participant in the COVID-19 response and whose duties require the use or access of vaccine administration data. Authorized users will be different for the different platforms since they each have different purposes and levels of data, as set out in the DUA and corresponding appendices. In addition, for each data platform, there may be Authorized Users who have incidental access to data in order to provide technical support services. For example, although CDC staff will not have access to the COVID-19 Data Clearinghouse, Oracle contractors may be designated as Authorized Users for technical support. For Authorized Users who may access data in order to support the COVID-19 response, through the creation of data analyses or aggregate data, their access would be limited to the IZ Data Lake.

   The Authorized User must be authenticated as a participant in the COVID-19 response by their parent agencies (e.g. CDC, HHS), and upon the termination of the user's time on the response, their access to data within data systems will be terminated. Examples of Authorized Users include, but are not limited to:
   - CDC COVID-19 FTE Response Staff
   - CDC contracted staff hired for specific duties and tasks related to the COVID-19 response
   - CDC Fellows who have been deployed to the COVID-19 response
   - HHS COVID-19 FTE and Contracted Response Staff

3. **Will jurisdictions be notified when there is an update or change to an appendix?**

The appendices have been included to provide detailed information about each of the data platforms and the data elements that may be received. Because some of the technical details about the platforms may change, these appendices will continue to be living documents that may need to be updated. Before any modifications to an appendix are made, jurisdictions will be notified and provided an opportunity to comment. If any data elements are modified, the jurisdiction will have to agree to and execute the change in data provision. CDC and HHS will not have the technical capacity to change the scope of the data unilaterally. The language of the DUA, including the limitations set out by state law and the general data sharing structure described in it, will remain the same. Therefore, any changes to data elements that may be reported will have to occur through mutual agreement.

CDC has sought minimally necessary data elements for the public health purposes set out in the DUA. Those elements are set out in the appendices. CDC will not seek social security numbers, driver's license numbers, or passport numbers. As stated here, any changes to the data elements will be in consultation with the jurisdictions.

4. **CDC mentions "Federal Partners" within the DUA. Can CDC limit who the federal partners would be?**

As set out in the purpose of the background and purpose of the DUA, the use of vaccine administration data will be limited to completing work in furtherance of the public health response to COVID-19. Since data may only be used in furtherance of the public health COVID-19 response, data about individual vaccine recipients may not be used to market commercial services to individual patients or nonpatients, to assist in bill collection services, or for any civil or criminal prosecution or enforcement, including, but not limited to, immigration enforcement, against such individuals whose information is shared pursuant to this DUA.

The response to COVID-19 is a whole government effort and may include agencies who are not typically involved in public health work. For example, the Department of Defense is part of Operation Warp Speed, the program under which the vaccine is being distributed. Authorized Users may need to access different data systems to provide technical support. However, the DUA is set out so that information that is pulled for data analysis or review can only come from the IZ Data Lake or HHS Tiberius. Thus, Authorized Users who are not providing technical assistance to any of the platforms may only have access to the redacted data that comes from the IZ Data Lake or Tiberius.

5. **Will CDC notify jurisdictions when publishing any of its data?**

Given the emergent nature of the response, HHS and CDC will make all efforts possible to notify the jurisdiction of the publication of data from the jurisdiction but may not be able to inform the jurisdiction in all instances. For such publications CDC will coordinate as soon as possible and practicable, with the fullest intent of protecting the data from the jurisdictions while functioning in the role of a data steward. As appropriate, publications will acknowledge the jurisdiction as the source of the data in any such publication. Also, consistent with federal law and CDC publication processes, CDC protects the identity of individuals whose data is being used in publications.

6. **Will jurisdictions receive notice prior to judicial or other legal, e.g. FOIA, disclosures?**

When there is a legal request, such as a FOIA request or judicial subpoena, for confidential information provided pursuant to this DUA, CDC program staff will provide timely notice to the appropriate jurisdiction.

7. What happens to data when the term of the DUA expires?

When the DUA expires, access to the data provided through the DUA expires with it; however, there are two circumstances that require the data to be maintained after the expiration of the DUA. First, due to the Federal Records Act, an archival copy of the data will need to be maintained by HHS and CDC. All protections, including limitations, for the data will survive the termination of the DUA as noted on pg. 8. Second, because there may be publications that rely on the data, the data must also be maintained for research integrity purposes. Data will be maintained according to Federal Records Act record scheduled (CDC Scientific and Research Project Records Control Schedule (N1-442-2009-1)). Should the veracity of a publication come under question, the data must still be available for CDC or HHS to defend the conclusions made with the data.