Freedman HEALTHCARE

BETTER DATA FOR BETTER SOLUTIONS

Releasing APCD Data: How States Balance Privacy and Utility

Freedman HealthCare, LLC 29 Crafts Street, Suite 470 Newton, MA 02458 **V** 617-243-9509 **t** @FreedmanHC

FreedmanHealthCare.com

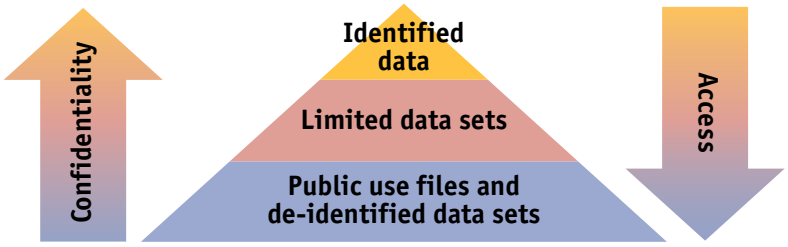
Which states currently release data to the public?

Currently, twelve states[1] allow non-state agencies to request APCD data: Arkansas[2], Connecticut, Colorado, Maine, Maryland, Massachusetts, Minnesota, New Hampshire, Oregon, Rhode Island, Utah, and Virginia.

How do states determine what data can be released? How do they develop data release procedures?

What and how data can be released depends upon the state's laws and regulations governing the APCD. States typically look to the HIPAA Privacy Rule for guidance, the framework for how health-care organizations may use and disclose patients' medical data, known as "protected health information."[3]

When new APCD states develop their data release models, other states' and CMS' data products and release processes offer useful guidance and structure. While each state has slightly different products and procedures for accessing them, states typically release one or more of the following types of data sets: de-identified or public use files, limited or comparable data sets, and/or identifiable information. De-identified and public use files have the least amount of member-level detail, and are easier to access, whereas identifiable data sets contain the most confidential information, and access is limited.



As confidentiality of data increases, access becomes more restricted.

States' data release products and processes vary depending on the type of data they are mandated to, or allowed to release according to their regulations and data release policies; whether release is limited to certain use cases; and pricing models.

For a detailed comparison of states' data products and processes, contact Alyssa Harrington at aharrington@freedmanhealthcare.com

[1]States included are those with legislation that establishes an APCD, and allows for public release of the data to external users. States that publish reports and analyses done in collaboration with the APCD, but that do release files for independent analyses, are not included.

[2]Arkansas is in the process of developing its data request process. For more information, contact Kenley Money at *KMoney@uams.edu* or 501-525-2244.

[3]Information about the HIPAA Privacy Rule and PHI can be found at *https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/#protected*

APCD Data Release Across States

	STATE APCD	CO	CT	ME	MD	MA	MN	NH	OR	RI	UT	VA
Allowable Release	Public reporting only						✔					
	Minimum necessary data	✔	✔	✔		✔			✔			
	Certain data elements never released	✔	✔	✔	✔			✔		✔		✔
	Some direct identifiers under limited circumstances	✔		✔		✔			✔		✔	
Allowable Uses	General health or health care improvement purposes	✔				✔	✔			✔		✔
	Limited to specific use cases		✔	✔	✔			✔	✔		✔	
	Must support state aims/benefit public	✔			✔	✔			✔			✔
Pricing	Free		✔				✔	✔				
	Depends on data requested			✔	✔	✔			✔	✔	✔	
	Depends on requesting entity			✔	✔	✔				✔		
	Determined based on request	✔										✔

This chart is based on publicly available information from public websites.

What is protected health information?

Protected Health Information or PHI, is defined by the HIPAA Privacy Rule as "individually identifiable health information" held or transmitted by a covered entity, or its business associate in any form or medium, whether electronic, on paper, or oral. This includes information related to:

- the individual's past, present, or future physical or mental health or condition;
- the provision of health care to the individual; or
- the past, present, or future payment for the provision of health care to the individual, and that identifies the individual or for which there is a reasonable basis to believe can be used to identify the individual. Protected health information includes many common identifiers (e.g., name, address, Social Security Number) when they can be associated with the health information listed above.

States vary in the amount and types of PHI collected by the APCD. For example, some states collect no PHI, while others collect member name, health plan beneficiary number, and street address.

How can a data set be de-identified?

Most states follow the HIPAA Privacy Rule's "de-identification standard."[4] This allows for the disclosure of health information if it has been de-identified using one of two methods described below. If a data set is de-identified using one of these methods, it no longer contains PHI, and is therefore no longer subject to Privacy Rule protections.

Expert Determination: Using statistical methods to de-identify data according to the following criteria:

- Data set is certified as de-identified by a person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable;
- Applying such principles and methods determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and
- Documents the methods and results of the analysis that justify such determination.

Safe Harbor: De-identifying data by removing the 18 identifiers listed below. The covered entity must also have no actual knowledge that the remaining information could be used alone or in combination with other information to identify the individual who is the subject of the information.

1. Names.
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP Code, and their equivalent geographical codes, except for the initial three digits of a ZIP Code if, according to the current publicly available data from the Bureau of the Census:
    a. The geographic unit formed by combining all ZIP Codes with the same three initial digits contains more than 20,000 people.
    b. The initial three digits of a ZIP Code for all such geographic units containing 20,000 or fewer people are changed to 000.
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older.
4. Telephone numbers.
5. Facsimile numbers.
6. Electronic mail addresses.
7. Social security numbers.
8. Medical record numbers.
9. Health plan beneficiary numbers.
10. Account numbers.
11. Certificate/license numbers.
12. Vehicle identifiers and serial numbers, including license plate numbers.
13. Device identifiers and serial numbers.
14. Web universal resource locators (URLs).
15. Internet protocol (IP) address numbers.
16. Biometric identifiers, including fingerprints and voiceprints.
17. Full-face photographic images and any comparable images.
18. Any other unique identifying number, characteristic, or code, unless otherwise permitted by the Privacy Rule for re-identification.

[4]Information about the de-identification standard can be found at *https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/#protected*

What is a Public Use File?

Public Use Files, or PUFs, contain de-identified data, and do not allow for linking members across claims, payers or years. They range from very highly aggregated data tables to claims-level extracts. States typically make public use files available upon request without a Data Use Agreement, and do not require review by a data release committee. Almost all states offer a public use file, or de-identified data set. Access to these data sets is less restrictive than for other data sets.

What is a Limited Data Set?

A limited data set contains some PHI, but excludes 16 "direct" identifiers (listed below). Unlike de-identified data sets, limited data sets may include more granular geographic information, such as member city and zip code, as well as elements of date and other numbers, characteristics, or codes not listed as direct identifiers. Under HIPAA, a data use agreement between the covered entity and data user is required in order to use and disclose this information.

1. Names.
2. Postal address information, other than town or city, state, and ZIP Code.
3. Telephone numbers.
4. Fax numbers.
5. Electronic mail addresses.
6. Social security numbers.
7. Medical record numbers.
8. Health plan beneficiary numbers.
9. Account numbers.
10. Certificate/license numbers.
11. Vehicle identifiers and serial numbers, including license plate numbers.
12. Device identifiers and serial numbers.
13. Web universal resource locators (URLs).
14. Internet protocol (IP) address numbers.
15. Biometric identifiers, including fingerprints and voiceprints.
16. Full-face photographic images and any comparable images.

Almost all states offer a limited or comparable data set. Access to these data sets usually requires review by a data release committee and approval by the APCD administrator.

What is typically included in a Data Use Agreement?

A Data Use Agreement (DUA) specifies the terms and conditions under which the user may handle, use and disclose APCD data. Most APCD DUAs have the following terms and conditions:

- Users may not attempt to re-identify individuals using the data;
- The data may only be used for the purposes specified in the application;
- The data may only be accessed by individuals specified in the application;
- The data may only be linked to other data sources as specified in the application, and may not be linked to other data sources for the purposes of re-identifying a member;
- Any publications, analyses or other outputs derived from the data cannot display small cells (typically 10 observations or less) to prevent identification of individuals. Some states require users to submit products derived from the data to the state prior to publication to ensure these conditions are met;
- The organization will implement security protocols and safeguards to ensure the data is protected at all times, and is only accessible to authorized individuals for authorized uses;
- Any security breaches or unauthorized uses of the data must be reported to the state within a certain timeframe. The organization may be required to implement a mitigation plan or destroy the data as a result;
- Upon the project's end, or as directed by the state, the organization must destroy the data in accordance with applicable state and federal laws around data destruction.

Do states usually release directly identifiable data?

Although many states do have a mechanism for releasing directly identifiable information from the APCD (data containing one or more of the 16 direct identifiers designated by the HIPAA Privacy Rule), users are carefully screened through a data request process. Some states limit who may access identifiable data based on prospective user categories, such as "academic researchers," "state agencies," and "commercial interests." Because of the sensitive nature of this data, uses of identifiable data are usually restricted to care coordination, treatment or other healthcare operations purposes.

How do states protect against misuse of APCD data?

States try to prevent misuse of APCD data through Data Use Agreements that limit uses to only those approved in the application, and that require certain security protocols. Some states can conduct an inspection to ensure compliance with the terms of the DUA. If a security breach occurs or is suspected, states may investigate the incident or require the user to destroy or return the data. Some state laws and regulations allow financial and/or criminal penalties for violations of APCD policies.

How do data request processes differ for different types of data sets?

Because they contain less detailed data and no identifiers, public use files generally have broader use cases, are less expensive, and have an expedited application and review process. Most states require only a data request form or brief application to request this type of data, and do not require review by a data release committee.

Limited data sets and other comparable claims-level data sets usually require a more extensive request and approval process, including a full application, a Data Use Agreement (DUA), review by a data release committee, and approval by the APCD authority. Some states, such as Maryland, also require review by an Institutional Review Board (IRB), an independent committee that reviews, approves, and monitors human subject research projects.

Requests for identifiable data are permitted for only very limited uses, and often restricted by type of user. These requests require a full application, DUA, and typically review by both a data release committee and another group, such as an IRB, Department of Justice, or privacy committee. These requests may also require additional agreements beyond a DUA.

How do states determine prices for APCD data?

States determine APCD pricing based on what is allowable according to their APCD rules and regulations; the cost of producing and transmitting files, maintaining the database, and/or providing user support; stakeholder feedback; and what other states charge for similar data. Pricing for APCD data products usually depends on the type of file requested, and the type of organization requesting it. Some states use a cost template to provide a customized quote for each request (Colorado and Virginia), whereas others have fixed fees (Massachusetts, Maine, and others). Most states charge fees for claims-level data sets, whereas some states, such as Minnesota and Connecticut, do not charge for highly aggregated public use files.

How are APCD data release products evolving?

As states' APCDs mature and data collection and aggregation processes become streamlined, states are looking to more innovative products to meet data users' emerging analytic needs. They are also looking for ways to leverage these products to help sustain APCD operational costs beyond state and federal funding for initial development. Examples of emerging products include customized, provider quality reporting; subscriptions to business intelligence tools that allow users to independently query databases; institutional licenses for academic research centers to use data for multiple projects; and integration of claims and clinical data for care coordination and treatment purposes. These products fit within APCD privacy parameters, while seeking to expand APCD use cases and user communities.

This report was produced by:
Alyssa Harrington, Consultant
Freedman HealthCare, LLC
aharrington@freedmanhealthcare.com

In collaboration with:
Linda Green, Vice President of Programs
Ildiko Kemp, Project Assistant

Acknowledgments

Freedman HealthCare would like to thank the following organizations for their collaboration in confirming data release products and procedures:

- Arkansas Insurance Department
- Colorado's Center for Improving Value in Health Care
- Maryland Health Care Commission
- Massachusetts' Center for Health Information and Analysis
- Minnesota Department of Health
- New Hampshire Department of Health and Human Services
- Oregon Health Authority, Office of Health Analytics
- Rhode Island Department of Health
- Utah Department of Health, Office of Health Care Statistics
- Virginia Health Information

Freedman HealthCare (FHC) is the leading national expert in APCDs, and has formally provided support to nearly 20 APCD teams nationwide. Our expertise includes feasibility studies and needs assessments, planning and infrastructure development, and project management and operations oversight. FHC specializes in APCD data release processes, products and pricing, and is experienced in helping APCD administrators develop and implement high quality, well-used and innovative data products.

Information in this report is current as of March 1, 2017.