2017

Report to the Legislature
_____

Health Care Cost, Quality, and Equity Data Atlas
Technical Feasibility Analysis

California Health and Human Services Agency
March 1, 2017

Acknowledgements

Table of Contents

Executive Summary

In 2014, the State of California received a federal design grant from the Centers for Medicare and Medicaid Services. The grant provided funds for a technical feasibility analysis of how California might consider developing a cost and quality database that promotes transparency and monitors trends in health care delivery. In 2016, pursuant to Senate Bill 1159 (Hernandez), Chapter 727, Statutes of 2016, the California Health and Human Services Agency began to explore options for the development of a California Health Care Cost, Quality, and Equity Data Atlas, also known as a Data Atlas. The intent of this research was to identify ways in which cost, quality, and equity data, when utilized together, can support the development of innovative approaches, services, and programs in both the private and public sectors. These data have the potential to inform the delivery of health care that is both cost-effective and responsive to the needs of consumers.

This report summarizes the research and provides policy makers with a conceptual framework of what is required to develop a Data Atlas. It assesses the efforts underway in other states, while analyzing the existing public and private California systems to determine their ability to submit data to a Data Atlas, or be utilized, in its implementation. The comparative analysis of models used in other states includes Colorado, Connecticut, Massachusetts, Minnesota, New York, Pennsylvania, Utah, and Washington. The analysis suggests that each state has used a different implementation approach, though there are a number of common themes and lessons learned among the states analyzed. Moreover, the report includes an assessment of both public and private systems that are currently collecting and analyzing health care data in California. The assessment finds that no single system has all the necessary capabilities; however, there are systems that could be used to initiate a phased implementation of a Data Atlas.

Furthermore, the report includes potential use cases that could guide the development and implementation of a Data Atlas. In conversations with stakeholders, it was evident that uses cases are important in defining the scope and utility of Data Atlas. The system capability analysis found that the California Department of Health Care Services, the California Department of Public Health, and Office of Statewide Health Planning and Development systems and datasets, when combined, could provide some initial use cases to narrow the development options. These datasets can provide meaningful information, which provide a framework for the use cases moving forward. Finally, the report summarizes the various governance structures used in states across the country—they vary from state to state, and each have strengths and weakness.

This research is intended to inform the design and implementation of a potential Data Atlas. It is important to note that policy makers and stakeholders will have to consider varying factors when looking to develop a Data Atlas. This includes necessary legislation, potential use cases, funding and sustainability approaches, and governance and data models. The analysis also suggests that there is no single strategy that can be deployed to successfully implement a Data Atlas. It is evident by the approaches used in other states that a combination of varying models and strategies will need to be used when developing a Data Atlas in California. This approach will not only address the states geographic and demographic diversity, but also the complexity of the health care delivery system.

Background and Context

In 2013, the State of California received a $2.7 million Round One State Innovation Model (SIM) Design Grant from the federal Centers for Medicare and Medicaid Services (CMS) to support the development of a State Health Care Innovation Plan. In 2014, the California Health and Human Services Agency (the Agency) published the *California State Health Care Innovation Plan[1]* (Innovation Plan), which formed the basis for the application of a three-year SIM Testing Grant. The Innovation Plan focused on initiatives designed to achieve savings in the short-term, but set a vision for changes that would advance the transformation of the delivery of health care over the long-term.

The Innovation Plan is rooted in the *Let's Get Healthy California Task Force Report[2]* (Report), which sets a vision for what California will look like in ten years if we commit to becoming the healthiest state in the nation. Moreover, the Report provides a framework for assessing Californians' health across the lifespan, with a focus on healthy beginnings, living well, and end-of-life. It also identified three areas that most profoundly affect the health and health care landscape: redesigning the health care delivery system, creating healthy communities and neighborhoods, and lowering the cost of care. Importantly, the report makes clear that eliminating health disparities is an over-arching goal.

The Innovation Plan addressed all three aspects of the Triple Aim – better health, better health care, and lower costs—and focused on three overarching goals: (1) reduce health care expenditures regionally and statewide; (2) increase value-based contracts that reward performance and reduce pure fee-for-service reimbursement; and (3) demonstrate significant progress on the Let's Get Health California indicators. The Innovation Plan also includes four initiatives and six building blocks. The four initiatives are designed to achieve savings in the short term as well as set in motion changes that will advance health system transformation over time. The goals of the building blocks are to enable California to track costs and quality across various systems of care, increase transparency, encourage competition, and promote continuous improvement.

In 2014, the State of California received a $3 million Round Two SIM Design Grant, after applying for, and not receiving, a $100 million Round Two SIM Testing Grant. This second design grant focused on six specific projects intended to further the goals and objectives of the Innovation Plan. This included a technical feasibility analysis of how California might consider developing a cost and quality database that promotes transparency and monitors trends in health care costs and performance. The purpose of this project was to identify approaches to plan, implement, and sustain such a database, while addressing the following objectives:

- Create a robust reporting system that promotes transparency and monitors trends in health care costs and performance;

---

[1] California Health and Human Services Agency (2014). *California State Health Care Innovation Plan*. Sacramento, California.
[2] California Health and Human Services Agency (2012). *Let's Get Healthy California Task Force Final Report*. Sacramento, California.

- Serve as a building block of the Innovation Plan including data-driven opportunities to advance the Innovation Plan, the Let's Get Healthy California (LGHC) goals, and the Triple Aim;
- Reduce reporting requirements and workload on the health care industry by leveraging the full range of public and private reporting systems;
- Collect, compile and analyze data from health plans and insurers about the cost, utilization and quality of public and private sector health care coverage and delivery of health care services; and
- Allow for the integration of other state health and human services data assets as appropriate.

The Agency contracted with Infiniti Consulting Group (Infiniti) to determine the technical feasibility of implementing a database. Infiniti interviewed stakeholders, assessed existing California data systems and reporting capabilities, and reviewed eight other states' cost and quality databases. In 2016, Senate Bill 1159 (Hernandez), Chapter 727, Statutes of 2016, codified this research, specifically to develop the California Health Care Cost, Quality, and Equity Data Atlas (Data Atlas), and the Agency began to research the options for developing a Data Atlas. The intent being to have cost, quality and equity data that encourages California health plans, insurers, and providers to develop innovative approaches, services, and programs. These data have the potential to inform the delivery of health care that is not only cost effective and responsive to the needs of enrollees, but also recognizes California' diversity and the impact of the social determinants of health. Combining the research, this technical feasibility analysis includes:

- A comparative analysis of models used in other states, including the various financing approaches and types of health care data utilized;
- Identification of key data submitters;
- An assessment of capability and opportunity for implementation; and
- An analysis of varying governance structures.

The stakeholder interviews were an informative and critical component of this report's analysis. Through these interviews, stakeholders shared insights on various issues, including governance models, sustainable financing mechanisms, data types (administrative data, claims data, encounter data, health equity and disparity data, and clinical data), privacy safeguards, data sources, and potential use cases. The stakeholder groups included:

- Data Submitters: Potential data submitters are health plans and insurers, and providers, including physicians and hospitals.

- Data Aggregators: Existing data aggregators are public and private entities located in California that currently collect data similar to that which could be collected by the Data Atlas.

- Consumer Advocates: Consumer Advocates represent people whose data would be submitted by data submitters to the Data Atlas.

- Subject Matter Experts: Although all interviewed participants could be considered a subject matter expert, the entities listed are those that did not fit into other categories, and/or were out of state entities.

Comparative Analysis of Potential Models

This report looks to assess the technical feasibility of a Data Atlas, which expands upon the concepts of a typical All-Payer Claims Database (APCD), but attempts to include more than just claims data to help understand and improve the quality and cost of health care delivery. For this report, we assessed eight states where a form of a Data Atlas was either fully implemented or in varying stages of implementation. The comparative analysis includes Colorado, Connecticut, Massachusetts, Minnesota, New York, Pennsylvania, Utah, and Washington. Appendix 1 includes a table that presents each state included in the detailed analysis along with key attributes identified for comparison. The analysis of these eight states revealed common themes for success, including:

- Legislation that mandates data submission from payers and other data submitters;
- Emphasis on defining specific use cases;
- Implementation using a phased approach;
- Early and ongoing stakeholder engagement ; and
- Establishment of governance, scope, privacy and confidentiality, and reporting requirements.

Figure 1: Comparative Analysis States

Colorado

In 2010, Colorado state law established the Colorado APCD, which contains 510 million health insurance claims and encounters from 21 commercial health insurance carriers, with products for Medicare, Medicare Advantage, and Health First Colorado (Colorado's Medicaid program). The Center for Improving Value in Health Care (CIVHC), a third party non-profit organization, appointed by the Department of Health Care Policy and Financing (HCPF), acts as the administrator of the Colorado APCD.

Governance Structure

HCPF establishes all rules associated with the Colorado APCD, including how to protect and release data. The Administrator ensures the APCD aligns with all Colorado statutes. Moreover, House Bill 10-1330 provides statutory authority for the HCPF to establish an Advisory Committee. The Advisory Committee makes recommendations on how to administer the database. Additionally, a Data Release Review Committee, established by the HCPF, is required to develop protocols for data release, review requests for Colorado APCD reports, and advise CIVHC on the appropriateness of those requests. The Administrator is required to make annual reports to the General Assembly and Governor. The Administrator initially worked with a data management vendor in early 2012 to begin collecting three years of historic claims information. In 2016, CIVHC announced planned improvements to the Colorado APCD including streamlined data intake, tracking, and processing.

Information Architecture and Rules

CIVHC published a data dictionary and data submission guide for payers to use to transmit data in a standardized format, while referencing industry data standards.[3] The State Administrative Rules, in the Code of Colorado Regulations[4] defines the payers that must submit claims data to the Colorado APCD as insurance carriers covering 1,000 or more lives, and offering health benefits plans, dental, vision, limited benefit health insurance, and short-term limited-duration health insurance. A "private health care payer" means a self-insured employer-sponsored health plan covering 100 or more enrolled lives in Colorado. Additionally, the rules exclude self-insured employer-sponsored health plans administered by a third-party administrator or administrative services-only organizations that service less than 1,000 enrolled lives in Colorado.

The data submission guide provides specific instructions regarding file submission methods, data file formats, data quality thresholds and the types of claim or encounter data submitted. Data files are submitted using Secure-FTP or the Administrator's Web Portal. The Administrator has also published data release rules, which mandate the requestor to demonstrate that the intended use supports the Triple Aim of better health, better care, and lower costs. In order to obtain data that contains Protected Health Information (PHI), requestors must submit a "finder's" file with information that can be matched to the data in the APCD. Only those records that are matched are provided back to the requestor. The requesting entity must enter into a Data Use Agreement (DUA) with CIVHC. CIVHC is required to report data requests annually to HCPF including how the data were used and

---

[3] Center for Improving Value in Health Care (2016). *Data Submission Guide*.
[4] Colorado Secretary of the State (2007) *Code of Colorado Regulations*, Department of Health Care Policy and Financing Rule 86-1.

how the data released met the Health Insurance Portability and Accountability Act (HIPAA) requirements.

New intake and verification processes improved the Administrator's ability to rapidly produce public data insights as well as custom data reports aimed at improving health and health care, while lowering costs. The Colorado APCD uses Microsoft SQL Server as the database engine and Microsoft Excel, Microsoft Power BI, and SAS as their data analytic tools.

Financing Mechanism

Capacity building grants from HCPF, The Colorado Trust, and the Colorado Health Foundation allowed CIVHC to develop and implement the Colorado APCD, with the expectation that it would become a self-sustainable resource. This funding supported the process of bringing health plan claims data on board and creating the infrastructure needed to develop custom reports and analytic tools. Necessary revenue to cover the ongoing costs of operating the APCD in 2016, and beyond, is expected from the sale of customized reports and data set fees.

Connecticut

In 2012, Connecticut state law established the Connecticut APCD, which contains medical claims and encounters, eligibility data, pharmacy claims, and provider data. Connecticut began collecting data from commercial health insurers and has plans to integrate Medicaid and Medicare claims in the near future. The Connecticut State Legislature passed legislation in 2015, requiring the State's Medicaid agency to begin providing Medicaid data to the APCD.

Governance Structure

Access Health Connecticut, Connecticut's health insurance exchange administers the Connecticut APCD. Access Health CT is overseen by a Board of Directors, which also approves policies and procedures for the APCD. The Board of Directors is comprised of 14 members representing a wide variety of stakeholder interests. The members of the board are charged with helping achieve the Exchange's goals. The APCD's Advisory Group provides guidance to the Board of Directors regarding strategic direction. The Advisory Group is made up of the commissioners (or their appointees) of the Department of Mental Health and Addiction Services, the Department of Insurance, the Department of Public Health, and the Department of Social Services. In addition to a health care advocate, the state's Chief Information Officer, the Secretary of the Office of Policy and Management, representatives of health insurance companies, hospitals, consumer advocates, experts of the health care industry, and health care providers serve on the Advisory Group. In 2015, Access Health CT signed a five-year contract with a data management and analytics services vendor, and by 2016 developed and integrated a web environment able to create population analytics and cost transparency reports.

Information Architecture and Rules

Access Health CT has developed and released a data submission guide, providing rules for entities required to submit data to the APCD. These "reporting entities", as defined in state statute, include: insurers licensed to do health insurance business in the state; health care centers; an insurer or health care center that provides coverage under Part C or Part D

of Title XVII of the Social Security Act; third-party administrators; pharmacy benefits managers; hospital service corporations; nonprofit medical service corporations; fraternal benefit societies that transact health insurance business in the state; dental plan organizations; preferred provider networks; and any other person that administers health care claims and payments pursuant to a contract or agreement or is required by statute to administer such claims and payments. A reporting entity not required to submit data are employee welfare benefit plans, as defined in the federal Employee Retirement Income Security Act of 1974.

In addition to the data submission guide, Access Health CT has also released a privacy and procedure policy, which includes rules pertaining to data release. The Data Release Committee reviews, approves, and denies data release applications submitted by applicants for the release of data. Any applicant must submit a complete data release application along with any applicable fees, be willing to be interviewed by the Data Release Committee, and enter into a DUA with Access Health CT.

Currently, the APCD is beginning the integration of dental, Medicare, and Medicaid data. By the end of 2017, Access Health CT's goal is to have a process in place for data distribution for research and in-house health care analysis. Additionally, Access Health CT hopes to release more advanced reporting on the website, and more sophisticated self-serving reports. In 2018, the intent is to implement a subscription model for provider and employer analytics to further drive the use of the data for population health management.

Financing Mechanism

To implement the APCD, the state applied, and was approved for $6,544,000 in federal funding as part of the Connecticut Health Insurance Exchange Level Two Grant. The Advisory Board is in the process of identifying a business model that will sustain the APCD into the future.

Massachusetts

In 2009, state regulations established the Massachusetts APCD, which collects medical (claims and encounters), pharmacy, and dental claims. The APCD also collects member eligibility, provider, and product files encompassing fully insured, self-insured, Medicare, and Medicaid data. This represents approximately 95 percent of all non-Medicare payers in the state. It excludes data from Workers' Compensation claims, TRICARE (the Military health plan), Veterans Health Administration, and the Federal Employees Health Benefits Plan. The Center for Health Information and Analysis (CHIA), an independent agency established by state statute in 2012 and serves as the Commonwealth's primary hub for health care data, administers the APCD.

Governance Structure

As a single independent agency, CHIA offers the benefit of having the infrastructure to work collaboratively with payers to improve data quality and completeness and the specialized staff needed to normalize data across payers to support cross-payer analyses. All requests for Massachusetts ACPD data are reviewed for compliance with CHIA's legal requirements. CHIA's Data Privacy Committee (DPC) reviews non-government applications, according to CHIA's statute and regulations. CHIA's Data Release Committee (DRC),

composed of a variety of external experts representing carriers, providers, researchers and consumers, provides additional scrutiny to certain applications to ensure that data release is in the public interest as defined by CHIA regulations.

Information Architecture and Rules

CHIA collects data monthly through a web-based transaction service. By using a web portal, payers are able to view their online submission logs to verify transfer status and where applicable, payers can download submissions to edit reports. Data, aggregated to the organizational level (e.g., cost/financial/statistical reports), do not require encryption software. However, CHIA provides payers with software that will encrypt and decrypt patient-level data. Data submission to the Massachusetts APCD is detailed in regulations, periodic administrative bulletins, and submission guides. Data submitters are required to complete and sign a data reporting agreement, and individuals that will be filing for the submitting entity must complete a DUA. The user agreements identify the data reporter's particular employees or agents/contractors and the particular submissions to which each require access. The Data Reporter can restrict the access of information to its employees by submission type.

CHIA releases data extracts to state agencies, payers, providers, provider organizations, and researchers. The Massachusetts APCD classifies data elements into three levels, based on their relative risk to patient privacy: Level 1- data are de-identified, per federal privacy law; Level 2- data elements include limited patient-level information, and therefore, pose a risk of re-identification; and Level 3- data elements contain direct patient identifiers (i.e. patient name, social security number, and date of birth). CHIA's Executive Director makes the final decision for release, based on the nonbinding recommendations of the committees. Before receiving the data, applicants must sign CHIA's DUA, which imposes requirements to maintain data security and protect patient privacy.

Financing Mechanism

The Massachusetts APCD has been funded, and is sustained by state general appropriations. Sustainability is supplemented by assessments on health insurers and hospitals; and fees from entities who wish to obtain data from the APCD. The fees are categorized by application, program support, and data use.

Minnesota

In 2009, Minnesota state law established the Minnesota APCD, which collects claims data from commercial payers, third party administrators/ self-funded payers, Medicare, and Medicaid. The Minnesota APCD emerged as an essential component of the health care reform package enacted by a bipartisan Minnesota State Legislature and signed by Governor Tim Pawlenty in 2008. The Minnesota Department of Health (MDH) administers the APCD, resulting in a state built data systems and data infrastructure. Moreover, it has also established analytic expertise to support a wide range of uses of the data. As of March 2016, the Minnesota APCD contains data on 4.3 million unique covered lives, roughly 89 percent of all Minnesotans with health coverage.

The MDH provides oversight and management of the data collection effort. MDH currently contracts with a vendor for services related to constructing and maintenance of the Minnesota APCD, including data collection, processing, quality assurance, and aggregation. In 2014, the Minnesota State Legislature directed the MDH to convene an advisory workgroup tasked with making recommendations to the Legislature about potential expanded use of the data, privacy and security considerations, as well as access and funding of the Minnesota APCD.

Information Architecture and Rules

As of July 2014, Minnesota statutes limit the use of the data to staff at MDH or organizations working under contract with MDH to conduct research on its behalf. In 2015, the Legislature directed MDH to produce annual summary information from the APCD that would be available for public use. Public Use Files provide the opportunity for researchers and the public to use the information contained in non-public datasets in an aggregated form that protects sensitive information. PHI is needed in order to create a de-identified member number and these data are encrypted. A detailed data dictionary has been released to data submitters informing them of which fields must be encrypted and which data standards are referenced by each data element. Each data element must follow the rules outlined in the Minnesota Data Submission Guide.

Health plans and third-party administrators are required to submit health care claims data for Minnesota residents if they paid at least $3 million in health care claims for Minnesota residents during the previous calendar year. Pharmacy benefits managers must submit health care claims data for Minnesota residents if they paid at least $300,000 in claims for Minnesota residents during the previous calendar year. Data submitters must use standardized reference codes when submitting member enrollment files, institutional and professional health care claims data, and pharmacy drug claims data. The data standards are aligned with nationally recognized standards that have been defined by Designated Standards Maintenance Organizations. Additionally, a data submitter must submit data elements to the data processor using the specifications outlined in the data submission guide. Any significant discrepancies in the data with respect to consistency, completeness, accuracy, or any other issue that may affect further review and verification of the accuracy of the data, must be reported to the data processor

Financing Mechanism

The Minnesota APCD is funded by the Minnesota General Fund. The state has invested roughly $3.2 million for the development of the APCD.

New York

In 2011, New York state law established the New York APCD. The New York APCD is in the design phase. The New York State Health Foundation, in collaboration with the APCD Council, has completed an in-depth stakeholder analysis that includes consumers, payers, providers, researchers, policymakers, and other APCD governance organizations. New York has developed a data intake solution, an interim APCD hosting solution, and a request for proposals for a data analytics vendor solution. This work has taken place alongside other State initiatives, including the development of New York's health information exchange

(Statewide Health Information Network for New York, or SHIN-NY) and health benefit exchange (NY State of Health Marketplace).

Governance Structure

The New York State Department of Health (NYSDOH) currently manages the New York APCD, which is also designed to serve the needs of the New York State Department of Financial Services (NYSDFS). The NYSDFS manages the regulation of commercial health insurance plans and intends to use the APCD to support rate review and other policy analysis. The APCD is being developed and administered through a new bureau within NYSDOH, the Office of Quality and Patient Safety. NYSDOH is also the agency responsible for the administration of New York's Medicaid program. Claims data will be integrated with clinical and quality data and public health repositories to create a more robust resource—the All Payer Database (APD). NYSDOH works closely with external stakeholders to gather input on the development of the APD. Representatives from health plans, provider organizations, researchers, and other interest groups participate on an APD Advisory Committee.

Information Architecture and Rules

NYSDOH's approach to APD development focuses on three core solutions: (1) data intake; (2) data warehousing; and (3) data analytics. In order to make immediate use of the data analytics capabilities, the data is being delivered from the Data Intake System into the Interim APD Data Warehouse (the OHIP DataMart). Currently, the OHIP DataMart receives data from the Data Intake System in five different models and houses New York's all payer hospital discharge data known as Statewide Planning and Research Cooperative System (SPARCS). The OHIP DataMart serves as a research and development environment for the receipt, processing, and analysis of new data sources.

The initial focus of the APD will be on claims and encounter data from health plans. The APD will add Medicaid and Medicare data, and integrate existing public health databases, laboratory data and clinical information from Regional Health Information Organizations (RHIO) and Electronic Health Records (EHR). The APD will also incorporate SPARCS data, which includes patient level detail on patient characteristics, diagnoses and treatments, services, and charges for each hospital inpatient stay and outpatient (ambulatory surgery, emergency department, and outpatient services) visit, and each ambulatory surgery and outpatient services visit to a hospital extension clinic and diagnostic and treatment center licensed to provide ambulatory surgery services. The APD will expand in subsequent phases and will have additional data sources added.

Integrating data from Medicare, (SPARCS), public health repositories, and the clinical and laboratory information from the Statewide Health Information Network for New York (SHIN-NY) will make it possible for the APD to provide more complete information integration for all participants in the health care system including policymakers, clinicians and consumers.

Financing Mechanism

The initial development of the New York APD is funded by a combination of state and federal funds. NYSDOH is exploring sustainable funding sources including: Medicaid federal

match funding, fees on health care providers and payers based on relative market share, general appropriations from the Legislature, fines collected from non-compliant data contributors, fees associated with the sale of data licenses, and/or providing de-identified data to users who pay a subscription fee.

Pennsylvania

In 2009, state statute established the Pennsylvania APCD. The Pennsylvania Health Care Cost Containment Council (PHC4), an independent state agency, administers the APCD. Pennsylvania currently collects hospital inpatient and laboratory data, outpatient and ambulatory data, eligibility and medical claims data. Pennsylvania is the only state that was analyzed that has an established process for collecting clinical data.

Governance Structure

As the administrator, PHC4 convened six committees and advisory groups, which include a wide cross-section of interests, and are composed of both government officials and representatives from stakeholder groups, such as consumers, hospitals, physicians, nurses, hospital quality improvement experts, health plans, commercial insurance carriers and health maintenance organizations.

Information and Architecture Rules

PHC4 collects data sets for inpatient hospital claims and ambulatory/outpatient claims, both based on the UB-04 claim form, laboratory data, cardiac surgery supplemental clinical data and payment data. PHC4 collects over 4.9 million inpatient hospital discharge and ambulatory/outpatient procedure records each year from hospitals and freestanding ambulatory surgery centers in Pennsylvania. PHC4 also collects data from managed care plans on a voluntary basis.

PHC4 publishes data submission guides on its website for each file type. Each submission guide has specifications for required data elements and required file format (fixed-width delimited flat text files). The guides specify record length, line delimiters, valid and invalid ASCII characters, and field justification rules. Data must be submitted using PHC4's secure portal. Additionally, the Pennsylvania reporting manual for insurers and third party payers has the layout and data submission instructions for member eligibility and medical claims data, which Pennsylvania began collecting January 1, 2010.

PHC4 has two different data request applications, depending on the entity classification:

- **Commercial:** Any organization or individual that requests data to resell or redistribute for a profit, including using the data in software products or using the data in analysis for profitable consultation with clients.
- **Non-Commercial, Research, and Government**
  - Non-Commercial: Any organization or individual that requests data to use internally for their own purposes and analysis.
  - Researcher: Any University-based researcher who requests PHC4 data for the benefit of the healthcare system.
  - Government: Any Pennsylvania State Agency that requests data to use for government-related purposes.

Both the data request applications require a DUA that ensures the applicant will not release any PHI data from the data sets or any data that could reasonably be expected to reveal the identity of any individual patient. Commercial applicants are prohibited from duplicating, selling, exchanging, publishing or otherwise releasing the raw, patient level data in any form. The government, research or non-commercial applicants may not further disclose any data other than for the purposes specifically identified in an approved application.

Financing Mechanism

The PHC4 activities are funded through the state's budget appropriations process. In addition, PHC4 receives revenue through the sale of its data to health care stakeholders.

Utah

In 2007, Utah state law established the Utah APCD. The Utah Department of Health Office of Health Care Statistics administers the Utah APCD, which contains data from health plans, Medicaid, and third party administrators in Utah. These data consist of medical, pharmacy, and dental claims, as well as insurance enrollment and health care provider data.

Governance Structure

The Utah Department of Health, Office of Health Care Statistics (OHCS) is currently responsible for building and managing the APCD. HB0009 amended the Health Data Authority Act to authorize the Health Data Committee, composed of 15 Governor-appointed members, to collect data on the costs of episodes of health care. Additionally, as funding is available, the legislation authorized the Department of Health to develop a plan to measure and compare costs of episodes of care

Information and Architecture Rules

Administrative Rules (R428-15-6) require commercial carriers licensed in Utah, with enrollment greater than 2,500, to submit data following the specifications published by the OHCS. Payers, also referred to as "carriers," submit flat files monthly, with some carriers having multiple reporting platforms. Utah developed a data submission guide for data submitters to provide member eligibility, medical claims and encounters, pharmacy claims and provider data. Each type of data has certain reporting requirements outlined in the data submission guide. The rule states that monthly files are required containing claim data for claims that were adjudicated in the previous month. OHCS has an online web application that documents requests for data access or release. The online application establishes the terms and conditions under which the Requesting Entity may use and disclose the data received through the request process. The Privacy Officer and the Director initially review all requests. Requests that require additional information or that involve identified data will require further review and can take up to a month to process. Requests that involve custom work are evaluated for feasibility, staff availability, anticipated cost, and the estimated time for completion. The application/ agreement allows OHCS to disclose certain data collected under the authority of the Utah Health Data Committee and maintained by OHCS to the Requesting Entity for purposes that are permitted by state law.

In September 2015, OHCS released user manuals for two types of Limited Data Sets that have been made available to private and public entities: 1) Claims Centric Limited Data

Set (CcLDS) and 2) Patient Centric Limited Data Set (PcLDS). CcLDS is designed to provide PHI to approved users that excludes certain identifiers but permits the use and disclosure of more identifiers than in a de-identified data set. The PcLDS is also a Limited Data Set that excludes certain identifiers but permits the use and disclosure of more identifiers than in a de-identified data set. It is important to note that unlike traditional limited data sets both data sets exclude the usually include service dates. Data collected by the Office of Health Care Statistics is available for research and statistical analysis by authorized users for approved purposes. Additionally, OHCS is currently working with the Transparency Advisory Group to select three new quality measures for early 2017 reporting. During the ongoing selection process, various health care partners have provided input.

Financing Mechanism

Initial funding and ongoing appropriations came from the State's general fund to finance and sustain the Utah APCD and provide start-up funding for the implementation of collection of clinical data. Utah receives annual, on-going funding of $615,000 from general funds and an $185,000 Medicaid match. The State also generates revenue by selling both limited data sets and claim level data sets.

Washington

In 2004, state statute established a voluntary APCD, hosted by a non-profit organization, Washington Health Alliance (formerly Puget Sound Health Alliance). The "Alliance" database initially focused on a five-county area in the Seattle region and contained commercial and Medicaid claim data. In 2014, state statute established the Washington APCD. The Office of Financial Management (OFM) oversees and manages the statutorily implemented statewide APCD. The first round of legislation only provided authority to collect Medicaid data. Amended in 2015, state law broadened the authority to collect commercial workers' compensation (Labor and Industries) and public employee claim data. Both efforts are running in parallel as of the date of this report.

Governance Structure

In 2016, OFM awarded the contract for an administrator to the Center for Health Systems Effectiveness (CHSE) at Oregon Health & Science University. CHSE is responsible for internal governance, management, determining sources of funding, and operations of the database. Moreover, CHSE will convene advisory committees for both data policy development and data release processes. The advisory committees must include in-state representation from key provider, hospital, public health, health maintenance organization, large and small private purchasers, consumer organizations, and the two largest carriers supplying claims data to the database. The governance structure and advisory committees must include representation of the third-party administrator of the uniform medical plan. A payer, health maintenance organization, or third-party administrator must be a data supplier to the APCD to be represented on the CHSE governance structure or advisory committees.

The statute specifies that CHSE must become certified as a qualified entity by the Centers for Medicare and Medicaid Services (CMS) in order to obtain Medicare claims data for analysis and integration with the APCD.

Information and Architecture Rules

The CHSE, in conjunction with OFM and the data vendor, will create and implement governance processes over the data, providing levels of access and use of the APCD. There are a number of mandated data submitters including: The State Medicaid program, public employees' benefits board programs, all health carriers operating in this state, all third-party administrators paying claims on behalf of health plans in this state, and the state labor and industries program. Development of the data submission guide will include processes for data submitters and other stakeholders to review drafts of the guide and provide input. OFM will have the final authority for approval of the guide. Additionally, the data vendor is required by statute to establish secure data submission procedures with data submitters and to maintain confidentiality of patient data, while in transit and when at rest, using state of the art encryption technology. They are also responsible for quality assurance, including verifying accuracy and validity of data submitted, consistency with specified file layouts, and identification of duplicate records.

The State is working on new sections of the Washington Administrative Code that will define rules for data requests and release procedures. Stakeholders can review the draft rules and provide feedback as to how the rules can best be structured. As part of the application process to receive and use data from the APCD, the requester will submit a data management plan, including: how data will be secured to maintain privacy and security, how data will be secured during electronic transmittal, and how data will be returned or destroyed following completion of the project. Claims or other data that include proprietary financial information, direct patient identifiers, indirect patient identifiers or unique identifiers may be released only to the extent such information is necessary to achieve the goals set forth in the statutes and with approval of an Institutional Review Board. Claims or other data that do not contain direct patient identifiers, but that may contain proprietary financial information, indirect patient identifiers or unique identifiers may be released to:

- Federal, state, and local government agencies upon receipt of a signed DUA with the office and CHSE;
- Any entity when functioning as the Lead Organization; and
- Government agencies that obtain claims data are prohibited from using such data in the purchase or procurement of health benefits for their employees.

Claims or other data that do not contain proprietary financial information, direct patient identifiers, or any combination thereof, but that may contain indirect patient identifiers, unique identifiers, or a combination thereof may be released to agencies, researchers, and other entities as approved by CHSE upon receipt of a signed data use agreement. Claims or other data that do not contain direct patient identifiers, indirect patient identifiers, proprietary financial information, or any combination thereof may be released upon request.

Reports utilizing data obtained through the formal request process may not contain proprietary financial information, direct patient identifiers or indirect patient identifiers. However, these data may be used to produce aggregate reports based upon geographic areas (with sufficient population sizes) gender, age, medical. The data vendor will establish a process for assigning unique identifiers to individuals, and protecting direct patient

identifiers and proprietary financial information. A signed data use and confidentiality agreement with CHSE must be delivered prior to release of these data elements. CHSE is also responsible for convening a Data Release Committee to review data requests and make recommendations for approval or denial. The State of Washington is considering using national data standards including standards that have been developed by the Accredited Standards Committee and the National Council for Prescription Drug Programs.

Financing Mechanism

Federal grants, awarded by the Centers for Medicare and Medicaid Services, provided the initial funding to implement the APCD. Currently, no grants cover ongoing operation and maintenance of the APCD. The enabling statute requires CHSE to be in charge of maintaining sustainability of the system. Additionally, it allows CHSE to establish fees for data products.

Capability and Opportunity Analysis

This section identifies a set of the current capabilities of public and private California-based systems, and evaluates the ability of those systems to submit data to a Data Atlas, or to be leveraged in its implementation. Information gathered during the stakeholder interviews has informed this capability analysis. The word "system" in this section can mean a single system or multiple systems operated and maintained by a single entity, and have capabilities in common with the implementation of a Data Atlas. Depending on the implementation approach, the systems outlined in this section will require further analysis to gather the necessary details required to move forward to support a Data Atlas. Additional criteria for future analyses include: the ease of use and efficiency of data input and retrieval, the maturity of the analytical tools and processes, security and encryption, and alternative means for accessing information and/or sharing information within and/or across systems. Further conversations with health plans and providers to assess their ability to submit data to a Data Atlas in a common format would be beneficial.

COMMON CAPABILITIES	
CAPABILITY TITLE	DESCRIPTION
Business Intelligence	Instruments are in place to transform raw data into meaningful and actionable information.
Data Extraction Standard	The ability of an entity to produce health care data files from an ANSI accredited organization: ASC X12, NCPDP, or HL7.
Data Quality	Data quality in this context includes completeness, comprehensiveness, and accuracy. The data reflects real-world events, collected in its entirety in a timely fashion, and processed using quality controls.
Data Submission Standard	The ability of an entity to accept health care data files from an American National Standards Institute (ANSI) accredited organization: ASC X12, NCPDP, or HL7.

COMMON CAPABILITIES	
CAPABILITY TITLE	DESCRIPTION
Interoperability	In this context, interoperability is an indication of how well a system securely exchanges data using defined interfaces or protocols with other systems.
Scalability	The ability of a technological system to increase its capacity to accommodate large amounts of data or to be able orchestrate data processing from multiple technological systems.
System Adaptability	A technological system (software and hardware) has the flexibility and ability to quickly adapt to change.

Public Data Systems

The preliminary system capability analysis finds the DHCS, CDPH and OSHPD systems and data sets provide meaningful information for a subset of the California population.

PUBLIC DATA SYSTEMS OVERVIEW		
	PUBLIC DATA SOURCE ENTITY	TYPE OF DATA
Member Data	**California Department of Health Care Services (DHCS)**	▪ Claim ▪ Encounter ▪ Provider ▪ Behavioral Health ▪ Enrollment ▪ Medicaid Care Coordination ▪ Child Health
	California Public Employees' Retirement System (CalPERS)	▪ Member ▪ Encounter ▪ Claim
	Covered California	▪ Member ▪ Provider ▪ Encounter (pending)
Population Data	**Office of Statewide Health Planning and Development (OSHPD)**	▪ Utilization ▪ Financial ▪ Inpatient Discharge ▪ Emergency Encounters ▪ Ambulatory Surgery ▪ Selected clinical ▪ Workforce
	California Department of Public Health (CDPH)	▪ Registry ▪ Utilization ▪ Provider ▪ Child Health

Department of Health Care Services (DHCS)

DHCS funds health care services for about 14 million Medi-Cal members. Nearly one-third of Californians receive health care services financed or organized by DHCS, making it the largest health care purchaser in California. DHCS operates and maintains multiple systems that in combination perform functions that are comparable to those that would be required as part of a Data Atlas. These systems include but are not limited to:

- Management Information System/Decision Support System (MIS/DSS)
- Post Adjudication Claims and Encounter Data Systems (PACES)
- Operating Rules for Information Exchange (CORE) interface or Secure File Transport Protocol (SFTP).

DHCS systems currently receive data from fee-for-service providers throughout the state, 26 contracted managed care plans, counties that provide various services, and other entities. After data is received using standardized data submission formats, the data is transformed and combined into a consolidated repository and analytic platform known as the MIS/DSS. DHCS has focused on stronger quality controls for processing encounter data, as well as exchanging data. Additionally, DHCS has a secure implementation of the CORE interface. The interface collects data from payers and other HIPAA entities. DHCS environment provides opportunities for lessons learned and may be considered, with respect to functionality, for the implementation of a Data Atlas. Claim level data may be shared for specific purposes but those purposes must also support the administration of the Medi-Cal program.

California Public Employees' Retirement System (CalPERS)

CalPERS serves more than 1.7 million members and administers benefits for nearly 1.4 million members and their families. Since 2003, CalPERS has processed encounter, pharmacy, lab, and claims data, making it an attractive analytical dataset. Although the data received is proprietary, CalPERS currently uses APCD defined format for data conversion efforts when it switched from Truven to Milliman. CalPERS is unique in that they capture actual cost data from their fee-for-service programs unlike other state data systems. However, the data cannot currently be shared outside of the organization. CalPERS would need to build a means to extract the data to submit to the Data Atlas if participation was mandated.

Covered California

Covered California (Covered CA) is the state's health insurance marketplace, and is the only marketplace where qualified individuals can receive financial assistance on a sliding scale to reduce premium costs. Covered CA securely transmits ASC X12[5] transactions bi-directionally with participating health plans. Their encounter analytic system is about to go live, at which time they will have the ability to leverage their business intelligence system. Covered CA provides detailed member information, but would have to build the functionality

---

[5] In 1979, the American National Standards Institute (ANSI) chartered the Accredited Standards Committee (ASC) X12 to develop uniform standards for interindustry electronic exchange of business transactions-electronic data interchange (EDI).

to produce related encounter data. Covered CA exchanges member information using standardized transactions with health plans, making it potentially easy to exchange the same data with the Data Atlas. The member data includes race, ethnicity, language, and other demographic details. Covered CA will soon be collecting encounter data from plans; however, claim level detail information cannot be shared outside the organization.

Office of Statewide Health Planning and Development (OSHPD)

OSHPD's systems have rich healthcare datasets that would be a beneficial resource in supporting the analytic goals of the Data Atlas. OSHPD systems collect a variety of datasets from licensed health care facilities including hospitals, long-term care facilities, primary care clinics, specialty clinics, hospices, and home health agencies, and make the data available to other organizations and the public in various formats. Data submissions are partially proprietary. However, there are plans to align submission standards and potentially use a combined set. OSHPD has the authority to collect administrative patient data under existing statutes and regulations. The data is used to inform a variety of use cases and analyses. OSHPD administers collected healthcare datasets in a standardized data warehouse and maintains analytical tools similar to those required of a Data Atlas. OSHPD also applies techniques to link patient data across datasets, which could prove to be valuable in the implementation of the Data Atlas.

California Department of Public Health (CDPH)

The Center for Health Statistics and Informatics (CHSI) at the California Department of Public Health maintains California's birth, death, fetal death, still birth, marriage and divorce records. CHSI coordinates department-wide initiatives to improve system effectiveness through better health information technology, increased epidemiologic analysis, strong liaisons with public health organizations and schools of public health, and effective partnerships with local health agencies and professionals. CHSI is working to create a culture of data sharing, which designed to promote better access to data to drive policy and program development. Moreover, the analyses done by the Office of Health Equity to identify and address the complexities of health and mental health inequities, while identifying interrelated and multisector strategies, are important and should be considered in the development and implementation of a Data Atlas.

Private Data Systems

In addition to the four public systems mentioned above, this report analyzes four different private data systems, which are defined as active California based health care data aggregators not governed by a public entity.

PRIVATE DATA SYSTEMS OVERVIEW	
PRIVATE DATA SOURCE ENTITY	TYPE OF DATA
California Healthcare Performance Information System (CHPI)	▪ Member ▪ Encounter ▪ Claim ▪ Pharmacy ▪ Health Care Quality

PRIVATE DATA SYSTEMS OVERVIEW	
PRIVATE DATA SOURCE ENTITY	**TYPE OF DATA**
California Integrated Data Exchange (CalINDEX)	▪ Encounter ▪ Claim ▪ Pharmacy ▪ Clinical
Integrated Healthcare Association (IHA)	▪ Encounter ▪ Claim ▪ Pharmacy ▪ Clinical
San Diego Health Connect (SDHC)	▪ Clinical ▪ Member ▪ Administrative (some)

California Healthcare Performance Information System (CHPI)

The California Healthcare Performance Information System (CHPI) is a 501(c)(4) nonprofit, public benefit corporation that collects proprietary formats of claim/encounter and eligibility data, but does not collect clinical data. CHPI collects claim/encounter and member data in proprietary formats from the commercial and Medicare HMO and PPO programs of Anthem Blue Cross, Blue Shield of California, and United Healthcare. Additionally, important to note is that CHPI is one of few Medicare Qualified Entities in the state of California. CHPI has a solid business intelligence foundation that continues to mature as data collection efforts improve. The submission of the data is not always timely and data consistency can be an issue as plans may have several warehouses, or data is lost or changed during claims payment processes. Data is submitted to CHPI on an annual basis, which needs to be taken into consideration if they are a submitter to a Data Atlas.

California Integrated Data Exchange (CalINDEX)

Anthem Blue Cross and Blue Shield of California collaborated in 2014 to implement Cal INDEX, an independent nonprofit organization that has developed a statewide Health Information Exchange (HIE). The organization will be merging with the Inland Empire Health Information Exchange (IEHIE), a move that will create one of the nation's most comprehensive nonprofit HIEs. The new HIE will combine the 11.7 million claims records from Cal INDEX founding members Blue Shield of California and Anthem Blue Cross with the 5 million clinical patient records of IEHIE and its 150 participating partners. Data is submitted in proprietary formats from the health plans. It has been recognized that Cal INDEX was challenging to implement, and as such, the data submission and extraction processes are still maturing. It is likely that, with time, the ability to submit data to a Data Atlas will improve.

Integrated Healthcare Association (IHA)

IHA is a nonprofit, nonpartisan group, which annually collects healthcare data in an aggregated and proprietary APCD format from participating plans. Due to the proprietary nature of their data collection, it would be difficult for them to provide data for the Atlas. However, IHA has a mature business intelligence program that can provide excellent examples for use cases of the Atlas. IHA has proven that progress is incremental and

improves over time, as is demonstrated with their,10 year running, Value Based Pay for Performance (VBP4P) effort.

San Diego Health Connect

SDHC, a nonprofit health information exchange, offers a diverse combination of capabilities due to the complexity of the health care environment in which they operate. They have a mix of security and transaction standards, built as a means to facilitate communication in the HIE landscape. As they do not store a wealth of data, having SDHC submit data to the Atlas could be challenging. SDHC was initially established as a centralized repository that distributed data to subscribers and implemented patient linking processes to formulate a master patient record for data exchange, which could be useful in the implementation of a Data Atlas. The SDHC system started as a federated system and transitioned to a hybrid model that stores pointers to data and systems. The system evolved over time to facilitate communication between entities. In addition, by storing some data in a central repository, the SDHC system was able to operate more efficiently.

Technology Alternatives

The report assesses three alternatives to developing a Data Atlas: Build New, Leverage Existing, and Network of Networks. Any future technology solution will be subject to the required information technology state approval processes and oversight. Moreover, funding for a state solution is subject to the annual budget process. A Data Atlas will need to be able to scale to handle an incredible amount of data. Any implementation direction will need to include detailed specifications of the capabilities of the system to accommodate growth as the system grows. There are many components that are combined to make a complete system, including software, hardware, security, network, storage, etc. that will need to grow with the system as the capacity of the system increases over time. Scalability for any system is often difficult to define, especially at the outset of an implementation, as the total capacity that a system will need to accommodate at maturity is difficult to define during the nascence of a project. Capacity will need to increase with system maturity, so each component of the system will need to scale equally with growth. The scalability of a system can be measured by improvements in processing when capacity is added, or when minimal degradation is realized as the demands on the system increase.

In developing a Data Atlas, it will be important to consider how it will manage the data within the system; this often is referred to as Master Data Management (MDM). MDM encompasses the efforts made by an organization to discover and define non-transactional lists of data, with the goal of compiling maintainable master lists. The result of a successful MDM solution is reliable, centralized data that can be analyzed and shared across enterprise systems, resulting in better business decisions. Examples of non-transactional lists that are candidates for MDM include a Master Patient Index, Master Provider Index or reference data sets such as ICD-9 and ICD-10 codes.

Build New

When building a new system (creating something from nothing), many barriers to system implementation are removed, such as integrating with legacy systems or leveraging existing network infrastructures. Building a new Data Atlas would still require integration

with data providers, but predefined formats will ease the burden of integration. Figure 2 depicts secure interfaces with data providers (left side) and consumers (right side), which would require implementation of the security standards outlined by HIPAA and HITECH at a minimum. There is also an additional security layer between the internal systems and the database layer, to further restrict access to the data and to augment security controls.

Figure 2: New Solution Build
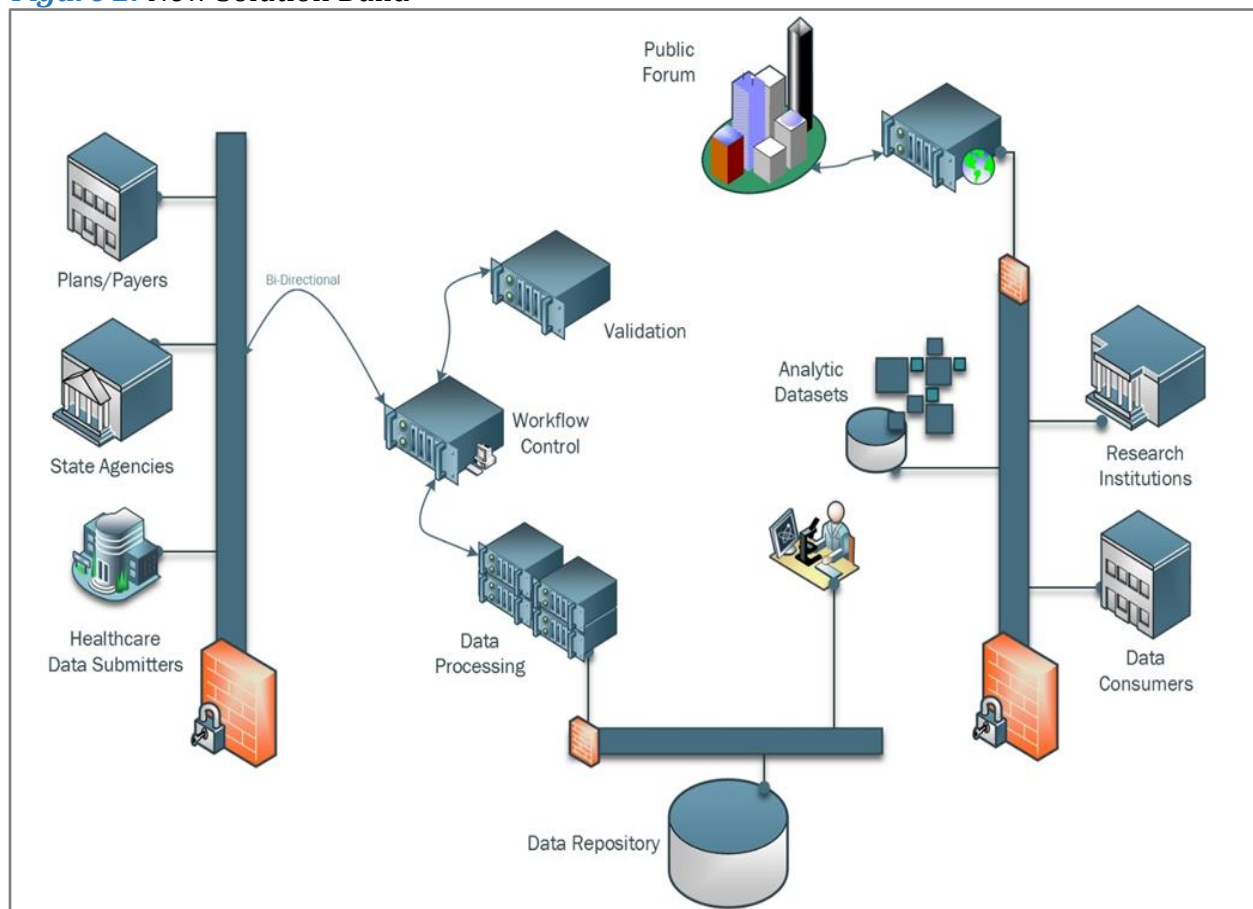
Building a modular system facilitates scalability and compartmentalizes system functionality, so components have dedicated responsibilities. The major areas that would be grouped, besides data, are workflow control, data processing, and data quality controls (validation). The workflow control component is essentially a traffic coordinator for the system; it acts as the communication control between other components of the system, and it manages the distribution of the computing resources. The workflow controller understands each component's key responsibilities regarding how data is processed, how integration between entities takes place, and how communication with other entities occurs.

The data processing component has a clear understanding of how to handle the various types of datasets that a Data Atlas would consume and generate. It would make decisions on how to load the data into a repository, manage data throughput, and scale resources to accommodate spikes in load. It would also provision analytic datasets to be consumed internally or externally. The validation component would have the responsibility of understanding the degrees to which data quality enforcement and controls would be

applied to datasets. It would have an understanding of all of the nuances of the previously mentioned standardized datasets and of non-standardized datasets that could be processed by a Data Atlas. The validation engine would be a business rule system whose domain is exclusive to data formats.

Leverage Existing

Leveraging an existing system to implement a Data Atlas would likely decrease the amount of time to address the goals in a defined use case. The leveraged system would need to be a large data processing system with some degree of modularity or flexibility. However, building on top of existing system does introduce challenges, such as potentially interfacing with legacy and inflexible systems. It also may be difficult to scale a leveraged system to meet the mature demands of a Data Atlas. If the legacy system does have modularity, then integration with new trading partners may be eased by using predefined formats for the new data streams, in which a new component could be introduced to facilitate processing.

Figure 3 depicts similar secure interfaces, like the new build, with data providers (left side) and consumers (right side), which would implement the security standards outlined by HIPAA and HITECH at a minimum. In the leveraged model, there would already be an existing secure interface that handles communication with existing data providers and consumers. There would also be an additional security layer between the internal (leveraged and new) system components and the database layer, to further restrict access to the data and augment security controls.

***Figure 3:*** Leveraged Solution Build
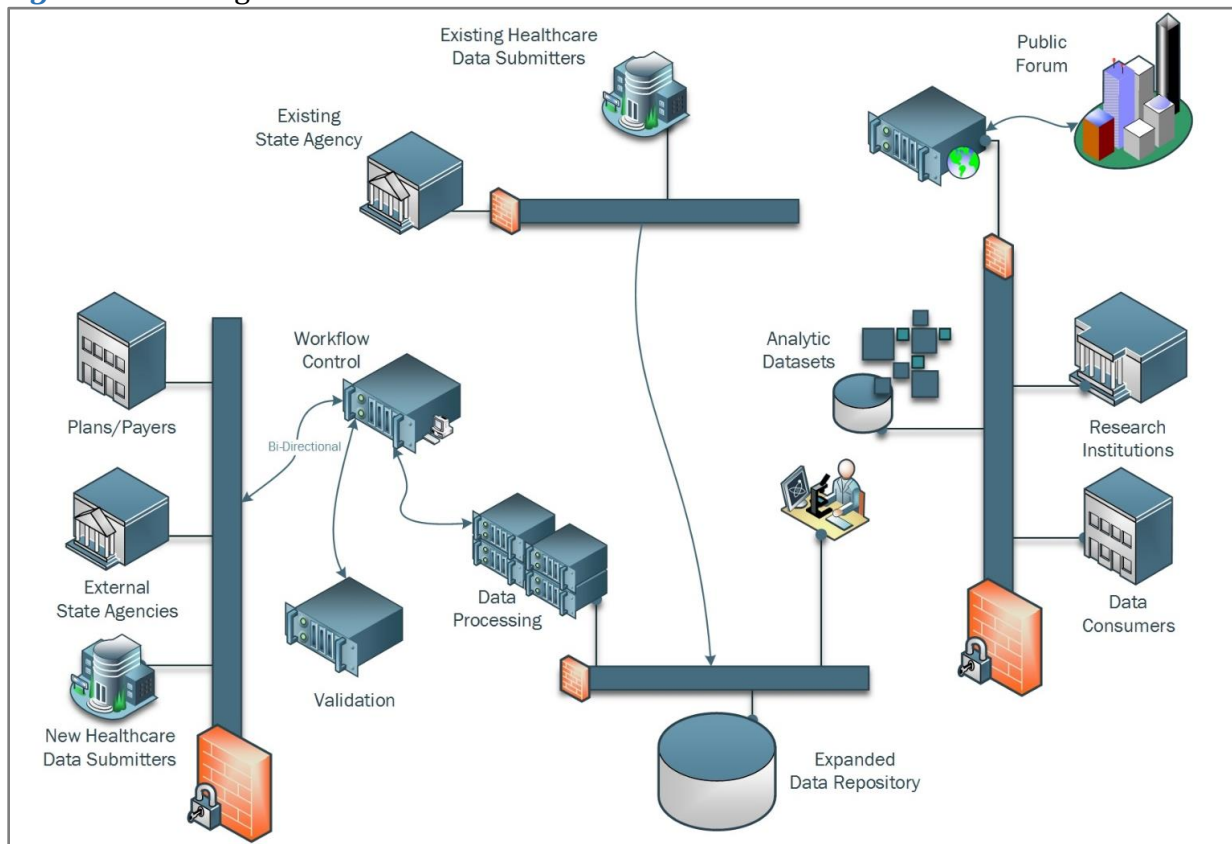
The leveraged system, like the new system, would be comprised of a variety of modular and loosely coupled components that would integrate with an existing legacy system to handle the processing requirements of a Data Atlas. If the structure of the legacy system allows for scalability, presupposing it is modular, then additional functionality could be added with new components, or the legacy system could be scaled to facilitate the additional load. Similar to the new system, the leveraged and enhanced system would have workflow control, data processing, and data quality controls (validation).

Network of Networks

Building a network of networks (NNs) is founded on the idea that a comprehensive network of health care data and systems can be linked together via existing systems or networks. The NNs would provide a framework of data collection that would allow for the seamless interchange of data between a Data Atlas and existing entities, rather than having one Data Atlas collect the data and store it in a single repository. The NNs would allow a Data Atlas to compile the results of other systems that already collect health care-related data, or that would collect it in the future. A Data Atlas would ideally also be able to discover new networks as they become members of the NNs.

Figure 4: Network of Networks Solution Build
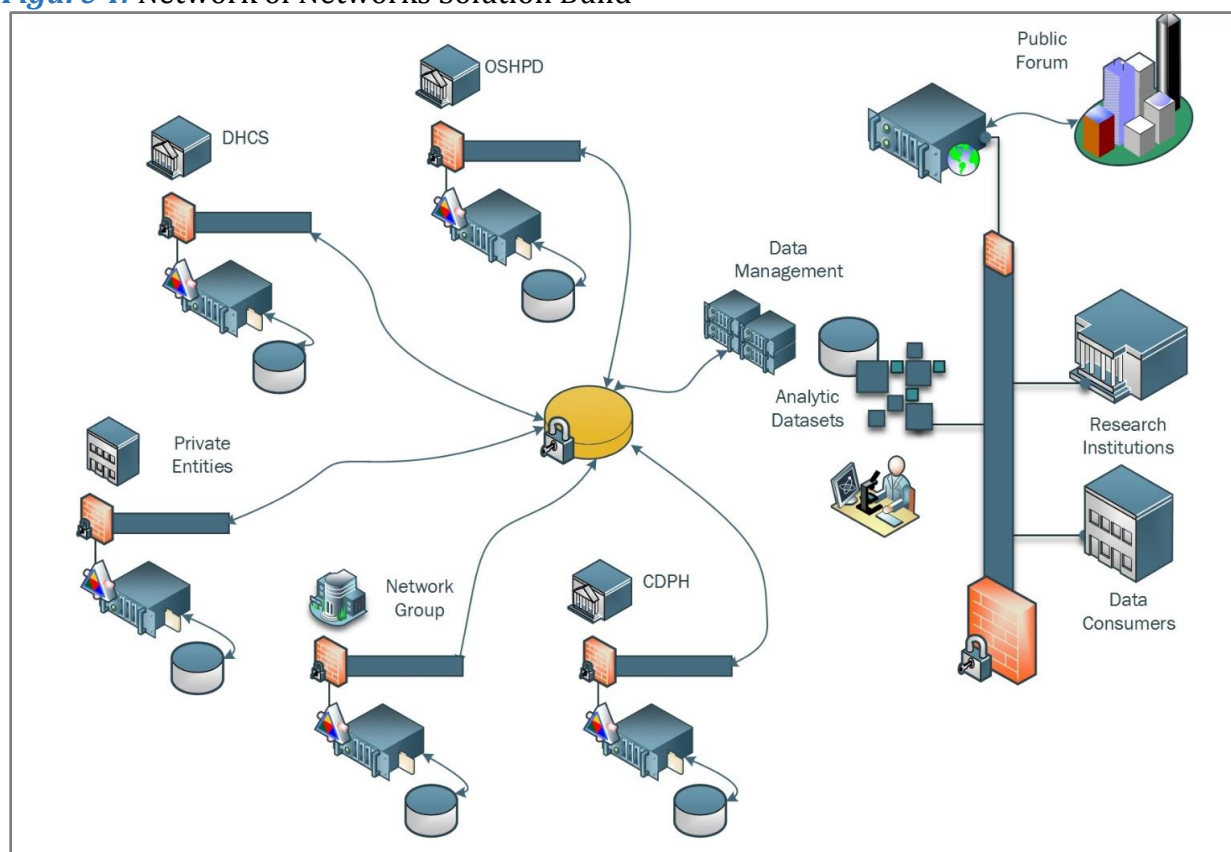
Building a NNs requires that all citizens of the NNs have a fundamental understanding of how to communicate with each other, which equates to knowing how to use Application Programming Interfaces (APIs). The APIs define the data exchange format and specifications so one data source knows how to connect to and exchange data with another data source (or node) on the network. Part of the communication includes security, of which, Figure 4

depicts similar secure interfaces as the new build, with each NNs citizen implementing their own security that would include the standards outlined by HIPAA and HITECH at a minimum. The NNs coordinator must also implement its own security in a similar fashion.

Potential Use Cases

There are many considerations and challenges inherent to establishing a Data Atlas. The elements of a database—technical requirements, governance structure, and funding sources—are interdependent. It was evident from our discussions with stakeholders and others states, that developing use cases was important in order to control the scope and ensure that a Data Atlas would answer important questions that would improve the delivery of health care. The potential use cases outlined below can be divided into three overarching themes: cost, quality, and equity- though many overlap between two of these categories. These are intended to help guide the development and implementation of a Data Atlas by narrowing the scope and achieving targeted outcomes.

Cost Use Cases
CalPERS Reference Based Pricing

In 2011, CalPERS began "reference pricing," which establishes a standard price for a medication, procedure or service and requires members to pay any charges beyond that price. The number of Anthem-CalPERS enrollees who chose a designated high-value hospital for their knee or hip replacement surgeries increased from 50 percent between 2008 and 2010 to 64 percent in the first nine months of 2012, compared with little to no change among Anthem policyholders not enrolled in CalPERS. Since its launch, the reference pricing initiative have expanded to include arthroscopy procedures, colonoscopies, and outpatient elective cataract surgeries.

Pharmacy Costs

A Data Atlas could potentially allow pharmacy cost to be assessed across the entire industry, for both public and private payers. There is a wealth of pharmacy data currently available in data systems within the purview of the Agency. Medi-Cal fee-for-service data provides the amount paid for each claim, while managed care plans also provide pharmacy level data. Since the vast majority of pharmacy data is reported on the National Council for Prescription Drug Programs (NCPDP) standard transaction, Extract-Transform-Load (ETL) needs could be minimized by a NoSQL solution that brings data together as needed, or with API.

Quality Use Cases
Maternity Care

Smart Care California is working with public and private payer organizations to utilize data to lower the rate of unnecessary C-Sections in the state. Increasing numbers of healthy women are undergoing obstetric procedures such as cesareans, repeat cesareans, and early elective deliveries when they may not be medically indicated practices that result in a higher rate of complications for mothers and babies. With more than 500,000 births every year in California, there is a compelling need to reduce unnecessary interventions and deliver appropriate, evidence-based care. Standardized performance measurement is needed to accelerate progress, increase transparency, and drive quality improvement.

Mental Health Impact on Physical Health

This use case requires the bringing together of the most disparate sets of data across the state. Health systems frequently do not house medical and mental health data within the same system and this is certainly true of Medi-Cal. A presentation provided during the Medicaid Enterprise Systems Conference opening plenary by the Missouri Medicaid Director discussed a striking example of how mental health issues affect physical health and mortality rates.

Equity Use Cases
Covered California Equity

Article 3 of the Covered California 2017 Individual Market Qualified Health Plan Issuer Contract, dated April 5, 2016, documents specific requirements for reporting on health disparities. The requirements include:

- Tracking and trending quality measures by racial or ethnic group, or both, and by gender for the Contractor's full book of business, excluding Medicare;
- Achievement of 80 percent self-identification by the end of 2019;
- Specific targets to be reported on for annual certification;
- Where self-reported information is not available, a standardized tool for proxy identification through the use of zip code and surname; and
- Contractor and Covered California will work together to assess the feasibility and impact of extending the disparity identification and improvement program over time. Areas for consideration include: 1) Income 2) Disability status 3) Sexual orientation 4) Gender identity 5) Limited English Proficiency (LEP).

Equity and Health Disparity

Disparities in outcomes by race and ethnicity, controlling for income, education, insurance status, could potentially be reduced by system interventions and changes in care. For example, a Data Atlas could be used to compare outcomes over time for cardiac care by race and ethnicity (controlling for socio-economic and insurance status).

There are a number of other potential use cases that could be explored. The following are some of the possibilities:

- Analyze how plan, payer, and provider relationships impacts fee-for-service usage in the marketplace, and how that influences capitation rates.
- Analyze regional variation to look at what type of association exists between health outcomes and quality of care when compared to costs.
- Compare costs and outcomes in regions that have a small amount of dominant health systems with regions that have a greater dispersion of providers. Track this comparison over time and look for any changes that occur as alignment increases.
- Examine the differences that occur between health systems that do and do not closely align their doctors, hospitals, and other elements.
- Look at the safety of care provided, by examining the number of medical errors, and the patient outcomes across both varying regions and varying health systems.
- Examine how health encounters are recorded and what can be done to improve standardization

Issues to Consider

When implementing a Data Atlas there are a number of issues to consider. Based on our comparative analysis of other states' efforts, it is clear there are a number of different approaches that California could consider. While, there is no one standardized model to follow, through the comparative analysis, some common themes emerged as best practices. These significant lessons learned from other states' successful implementations include:

- Establishing in legislation these elements:
    - Governing body and oversight
    - Scope of the data collection effort
    - Mandatory data submission
    - Privacy and confidentiality
    - General funding considerations
    - Reporting requirements
- Defining specific use cases
- Implementing the Data Atlas using a phased approach
- Engaging stakeholders early and on an ongoing basis
- Standardizing data sets
- Establishing an inclusive governance structure
- Exploring many options for funding and sustainability

Additionally, as the research for this report progressed, interviews with the stakeholders revealed that a technical solution to this database would not be easy, and there are many unknowns and undecided factors. In order to begin to bridge the gap towards successful implementation there is much more analysis required and decisions that need to be made before any technological solution is selected. Moving forward the following sections review the issues to be considered for further investigation.

Staffing Resources

If California does choose to proceed with the development and implementation of a Data Atlas, it may also consider dedicated staff resources. Staff could be used to evaluate the policy and technical requirements needed to successfully implement a Data Atlas. It is evident from the research that this is a complex endeavor that will require various skillsets and extensive expertise in health information technology and data management. Moreover, a combination of private and public sector experience may be necessary in order to effectively develop a solution that will not only meet the needs of the State, but would also align with the efforts underway in the private sector. Any funding for such resources would need to be subject to the annual budget process.

Use Cases

The Data Atlas should consider defining use cases to help narrow development options. To start, it is beneficial to develop a scoring mechanism with evaluation factors scored with a weighted value based on the response. The following table below provides a simplified example of how a use case could be evaluated with suggested questions and scoring. Scoring values in this example are 1 – 5, with 1 representing the least in agreement with the question and 5 the most in agreement. Using a purely numeric score might not be the main decision factor, depending on the situation.

These are just a few samples of questions that could be used in determining the priority of use case selection. Using a purely numeric score might not be the main decision factor, depending on the situation. Timing, for instance, could be more important in deciding to move forward with a certain use case and therefore could override other decision points. Likewise, ease of obtaining data could be the path of least resistance and so a decision could be made to choose a use case that could be fulfilled with currently available data sets. Furthermore, prioritization of identified use cases should be considered when thinking about how to implement a phased approach.

EVALUATION FACTOR NUMBER	USE CASE EVALUATION FACTOR	EVALUATION SCORE
1.	Does this use case provide a potential quick win that will show early value of a Data Atlas?	1
2.	Could the use case benefit a majority of stakeholders?	4
3.	Is the data needed for the use case easily available?	2
4.	Could the results of the use case contribute to funding and sustainability resources (e.g., special interest grant funding, federal matching funds)?	4
5.	Could existing systems or resources be leveraged to respond to this use case?	4
6.	Have any other APCD states produced results for this or a similar use case?	1

Legislation and Policy

If a Data Atlas were to be developed, California would need to create further legislation to establish the Data Atlas. The APCD Council has developed a model to assist states with development of legislation.[6] The model is based on a review of fifteen states' existing legislation and analysis of other federal and state regulations regarding data collection and release. The model is designed as a framework with guidance for states with considerations to tailor the legislation to each state's specific environment. It discusses the pros and cons of legislation versus rules for subjects such as data submission and release requirements. The legislation would need to describe the purpose and intent of the Data Atlas; establish the governance structure; define the scope, including the data submitters; describe how privacy and integrity of the data will be protected; describe the intended funding mechanism; and outline public reporting requirements. Moreover, it will be important to consider federal regulations and standards that govern data collection, at the U.S. Department of Labor and the U.S. Department of Health and Human Services.

---

[6] Porter, J., Love, D., Costello, A., Peters, A. and Rudolph, B., *All-Payer Claims Database Development Manual: Establishing a Foundation for Health Care Transparency and Informed Decision Making*, APCD Council and West Health Policy Center, February 2015.

Governance Models

There are multiple types and levels of Data Atlas governance. One consideration is what type of high-level governance structure, responsible for setting policy, managing communications, directing operations and enforcing provisions set in statute, does the state want to implement. Additional considerations need to be given to the governance structure of the technical solution, data release and data submission. Regardless of model, stakeholders recommended that a board made up of public and private payers and include a technical advisory committee manage governance structure. The public and private payers should come from state agencies like DHCS, Covered California, and CalPERS. Private board members should include employer organizations, advocate groups, health professionals, and IT professionals with a background in health care In terms of the managing governance structure there are four potential models examined:

GOVERNANCE STRUCTURE TYPES	
GOVERNANCE MODEL	**DESCRIPTION**
Third Party Non-Profit	Third Party Non-Profit as required by legislation: In this model, the solution is hosted by a vendor, the state appointed a third party administrator, receives guidance from a statewide multi-stakeholder advisory group and follows the rules set by the state entity
	Non-Profit with a board of directors: In this model, the solutions are software as a service (SAAS) powered by two health care analytic firms managed by a non-profit organization.
Public Entity	Where the database is housed within a government department and managed by a consortium of government entities and stakeholders
Independent Government Agency	A stand-alone agency that reports directly to the Governor and collaborates with stakeholders
Public Private Partnership	A collaborative formed between a public entity and a third party non-profit to form a "governance organization."

Funding and Sustainability

After defining use cases, the issue of how to fund the development and implementation of a Data Atlas is important to consider for the future sustainability of a Data Atlas. Stakeholders suggested exploring federal match funding options, related, for instance, to the Medicaid Information Technology Architecture (MITA). Many states have used grant funding for both implementation and, in some cases, sustainability. The concept of data submitter fees could also be explored. A Data Atlas could select use cases that maximize federal funding while those opportunities are available. In addition, a fee schedule for those requesting data could be established as a measure of sustainability funding. A Data Atlas may

want to consider a subscription service as a sustainability source as well, where interested parties (e.g., payers, providers, or other healthcare organizations) pay a regular fee for access to the data. Funding for a Data Atlas would be subject to the approval of resources through the annual budget process.

Data Consideration

The states analyzed all have similar procedures surrounding the release of data to researchers or other government agencies. The states that do release data to non-government entities follow the Code of Federal Regulations and HIPAA rules and will release data only after the entity requesting the data completes an application explaining how the data will be used, and a DUA is executed.

In order to obtain Medicare data for California members, a Data Atlas would need to become a Medicare Qualified Entity (QE). To become a QE, the CMS Qualified Entity Program must certify the organization. Once an organization becomes a QE, they can receive Medicare claims data under parts A, B and D and use the data to evaluate provider performance. Additionally, QEs are required to produce and publicly share CMS- approved reports demonstrating provider performance. Lastly, as a QE the organizations are able to provide or sell data to authorized users. There are currently five QEs reporting on California, the Health Care Cost Institute (HCCI), Amino, OPTUMlabs, FAIR Health, and CHPI.

Since California is a predominately managed care state, some stakeholders have expressed concern regarding cost transparency for managed care capitated encounters. While a fee-for-service equivalent may be difficult to establish initially, it is possible to link Medi-Cal encounter data to the Medi-Cal capitation payment data to begin to get a picture of cost. Both sets of data are stored at the individually identifiable level, allowing for linkage from services to payments. This is a potential use case that could be satisfied by leveraging existing systems and data.

Privacy Considerations

In order to appropriately de-duplicate individuals and track people over time in a Data Atlas, PHI must be used. Depending on the model chosen, the exchange and storage of PHI may be required to achieve the goals of a Data Atlas. PHI must be managed to protect confidentiality of patient data. Federal regulations require that recipients of data sets containing PHI enter into a written DUA with the entity responsible for administration of the Data Atlas. A covered entity (in this context the Data Atlas administrator) may always use or disclose for research purposes health information which has been de-identified, meaning it is not individually identifiable and the covered entity has no reasonable basis to believe it can be used to identify an individual. Currently there are two ways to de-identify data as described in the HIPAA standard: (1) expert determination and (2) safe harbor. Expert determination is the process of applying certain principles and methods so 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. Safe harbor ensures that a series of eighteen identifies that are linked to the individual, relative, employer, or household member are removed from the data.

The level of depth of, and in some cases the staff involved in, the interviews did not fully yield visibility to the technologies used to establish these types of databases in other states or in the regional databases within California. Research could further the understanding of the technical details and newer technologies, for instance cloud solutions, or database applications that can bring together disparate sources of data without extensive extract transform and load processes needed to convert the data to a common format. These types of solutions cannot only decrease the time to production, they also can provide opportunities for cost savings to implement.

Change Management

Typically, causes of a project's failure may include insufficient executive support, inadequate funding, incapable software, or lack of interoperability of components. In reality, however, a project's success relies just as much on the software and hardware needs, as the time needed to transition stakeholders and staff to changing business processes. Given these factors, an organizational change management plan can create the process, tools, and techniques to manage the people-side of change to achieve the required business results. Change management focuses thinking from the As-Is to the To-Be state of an initiative, and emphasizes skill building, acceptance, resistance management, and communication. One of the most recognized is Prosci's ADKAR (Awareness, Desire, Knowledge, Acceptance and Reinforcement) methodology. The methodology provides guidance for defining the change, developing tools and templates for change assessment, accounts for roles (e.g., sponsors, change agents and stakeholders and direction for developing the organizational change management plan).

Conclusion

If California were to implement a Data Atlas, there would still remain a number of issues to consider. This research is intended to inform the potential design and implementation of a Data Atlas. It is important to note that varying factors will need to be considered when looking to develop a Data Atlas. This report identified best practices implemented by other states, as well as an analysis of California's existing capabilities. Any technology solution will be subject to the required information technology state approval processes, and funding for such a state solution will be subject to the annual budget process. Moving forward, this report can be used to initiate a conversation among policy makers and stakeholders if and how to proceed with the development and implementation of a Data Atlas.

Appendix 1: State Comparison Table

STATE	DEVELOPMENT PHASE	GOVERNANCE STRUCTURE	DATA SUBMITTERS AND RELATED DATA		DATA STANDARDS	DATA RELEASE
COLORADO	Operational	Government Agency - CO Department of Health Care Policy and Finance (HCPF), State Medicaid Agency HCPF appointed the Non-Profit Organization - Center for Improving Value in Health Care (CIVHC) as the Administrator of the APCD Data Vendor - Human Services Research Institute & Subcontractor National Opinion Research Center at the University of Chicago	Medical Claims	X	Data Submission Guide aligned with ANSI X12 HIPAA transaction sets (837, 835, 271) & NCPDP	CIVHC Data Release Review Committee (DRRC) Application Required: State agency or private entity engaged in efforts to improve health care or public health outcomes for Colorado residents may request data, including PHI. Data Use Agreement (DUA)
			Encounters	X		
			Pharmacy	X		
			Dental	X		
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity: i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health			
			Laboratory			
			Other			
CONNECTICUT	Operational	**Government Agency:** State Health Insurance Exchange, Access Health CT Board of Directors Advisory Group	Medical Claims	X	Data Submission Guide aligned with the ANSI X12 HIPAA transaction sets (837, 835) & NCPDP	Application, Data Review Committee and Data Use Agreement (DUA) Required
			Encounters	X		
			Pharmacy	X		
			Dental			
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare			
			Self-Insured			
			Equity, i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health	X		
			Laboratory			
			Other			

STATE	DEVELOPMENT PHASE	GOVERNANCE STRUCTURE	DATA SUBMITTERS AND RELATED DATA		DATA STANDARDS	DATA RELEASE
MASSACHUSETTS	Operational	**Government Agency** - Center for Health Information and Analysis (CHIA). CHIA is an independent state agency established through statute. CHIA manages and administers the MA APCD and has a dedicated staff for the APCD.	Medical Claims	X	Data Submission Guide alignment to national standards	Applications and Data Use Agreements (DUA) for: ▪ Government Agencies ▪ Non-Government Agencies Applications for data reviewed by CHIA's Chief Privacy Officer Data Release Review Committee (DRRC)
			Encounters	X		
			Pharmacy	X		
			Dental	X		
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity, i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health			
			Laboratory			
			Other - Product File & Risk Adjustment	X		
MINNESOTA	Operational	**Government Agency** – MN Department of Health **Data Vendor** - Onpoint Health Data	Medical Claims	X	Data Submission Guide aligned with national standards including ANSI X12 HIPAA transaction sets (270, 271, 835, 837), UB-04, CMS-1500 and NCPDP.	Use of data is restricted by statute to the Minnesota Department of Health, and to contractors working on the Department's behalf.
			Encounters	X		
			Pharmacy	X		
			Dental			
			Provider			
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured			
			Equity, i.e. Race/Ethnicity			
			Clinical			
			Behavioral Health			
			Laboratory			

State	Development Phase	Governance Structure	Data Submitters and Related Data		Data Standards	Data Release
			Other			
New York	Design	**Government Agency** – NY Department of Health (State Medicaid Agency) **Data Vendor** – Optum Government Solutions	Medical Claims	X	Data Submission Guide (in development) ANSI X12 Post Adjudicated Claims Data Reporting (PACDR 298, 299, 300) and NCPDP Post Adjudicated Claims	Will require an Application and Data Use Agreement (DUA)
			Encounters	X		
			Pharmacy	X		
			Dental	X		
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity, i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health			
			Laboratory			
			Other			
Pennsylvania	Operational	**Government Agency:** Pennsylvania Health Care Cost Containment Council (PHC4)	Medical Claims	X	Data Submission Guides in alignment with the Uniform Claims and Billing (UB-04) and ANSI X12 837, 835, 270, 271 transaction sets	Application and Data Use Agreements for ▪ Commercial Entities ▪ Non-Commercial, Government or Research
			Encounters	X		
			Pharmacy			
			Dental			
			Provider			
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity, i.e. Race/Ethnicity	X		
			Clinical: Cardiac Surgery Supplemental Clinical Data	X		
			Behavioral Health			
			Laboratory	X		

State	Development Phase	Governance Structure	Data Submitters and Related Data		Data Standards	Data Release
			Other Hospital Inpatient Ambulatory/ Outpatient	X		
Utah	Operational	**Government Agency:** Utah Department of Health, Office of Healthcare Statistics	Medical Claim	X	Data Submission Guide aligned with the ANSI X12 HIPAA transaction sets (271, 837, 835) & NCPDP	Application requires entity to specify their organization type as: ▪ Public Agency or Non-Profit Organization ▪ Profit/Private Sector Agency ▪ Data Supplier Data Use Agreement (DUA)
			Encounters	X		
			Pharmacy	X		
			Dental	X		
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity, i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health			
			Laboratory			
			Other			
Washington	Design	**Government Agency** - Office of Financial Management **Non-Profit Organization -** Center for Health Systems Effectiveness at Oregon Health & Science University is the Lead Organization selected as Administrator **Data Vendor -** Onpoint Health Data is the data vendor subcontracting with	Medical Claims	X	Under consideration ▪ Data Submission Guide aligned with ANSI X12, NCPCP & other national standards ▪ ANSI X12 Post Adjudicated Claims Data Reporting (PACDR)	Will require an application and Data Use Agreement (DUA)
			Encounters	X		
			Pharmacy	X		
			Dental	X		
			Provider	X		
			Eligibility	X		
			Commercial Plans	X		
			Medicaid	X		
			Medicare	X		
			Self-Insured	X		
			Equity, i.e. Race/Ethnicity	X		
			Clinical			
			Behavioral Health			
			Laboratory			

STATE	DEVELOPMENT PHASE	GOVERNANCE STRUCTURE	DATA SUBMITTERS AND RELATED DATA		DATA STANDARDS	DATA RELEASE
		the Lead Organization	Other Public Employees & Workers Comp.	X	▪ NCPDP post adjudicated claims reporting standard	