Kentucky Health Data Trust Initiative

APCD Infrastructure – Key Lessons Learned

Prepared for the Kentucky Health Data Trust
Interagency Governance Workgroup

Freedman HealthCare
June 4, 2015

Deliverable 5.3

INTRODUCTION

All Payer Claims Databases (APCDs) have been operating for over a decade in a variety of models and formats across the country. Each state's approach reflects a unique mix of policy goals, stakeholder engagement, privacy considerations, reporting requirements and political support.

The purpose of this document is to present the key elements encountered in building the technical infrastructure of an APCD. These are the topics that Freedman HealthCare (FHC) is most frequently asked to discuss when supporting states in APCD planning and development. FHC expects that more issues will emerge throughout the development of the Kentucky Health Data Trust (KyHDT), and FHC will offer additional information from other states as needed. No other state has integrated its APCD into a data warehouse that is interoperable with other state-agency human services data files. Furthermore, as of this writing, CHFS is in the process of prioritizing the particular use cases (reports and analysis) envisioned for the Data Trust. Therefore, these APCD lessons learned may or may not be relevant to the Data Trust.

LESSONS LEARNED

1. **Set expectations: APCDs in other states are highly useful for retrospective reporting, analysis, policy development and population health monitoring. Current data collection models are less appropriate for case management, care coordination and clinical outreach.**

APCDs typically collect data from carriers and public payers on a monthly or quarterly basis with quarterly refreshes of reports and data marts. The time between a date of service and viewing that record in the data mart could be as long as 11 months. Based on commercial data, approximately 95% of claims dollars are processed within 90 days of claims submission. Carriers require at least one month to prepare file extracts; the data manager's ETL and data mart refresh processes require up to 90 additional days. See Table 1 for an example of how the timeline for a quarterly submission and quarterly data mart refresh schedule might look.

Table 1: Example of Mandatory Quarterly Submission and Quarterly Refresh

Dates of Service	Claims Run out	Carrier File Submission Deadline	ETL, Processing, Data Mart Refresh (90 days*)	Reports Available
January through March	June 30	August 15	November	December
April through June	September 30	November 15	February	March
July through September	December 31	February 15	May	June
October through December	March 31	May 15	August	September

*This time period becomes shorter as the process becomes routinized.

Because of the multiple steps that health care claims must pass through from the date of service to the end of production in the APCD, there is usually a significant lag time for APCD data. Current data submissions are typically monthly or quarterly for most states, with the data warehouse usually

refreshed every three to six months. However, this timeframe can lead to outdated claims data. There is increasing interest in monthly updates to the data warehouse, particularly in Colorado and Massachusetts, as more frequent refreshes will provide much more timely data. However, no state has moved to monthly APCD refreshes to date, due to the cost and competing organizational priorities of multiple production cycles. States developing new APCDs should weigh these issues and determine the most feasible refresh schedule, and then set clear expectations among APCD stakeholders.

2. **APCD teams should develop a written set of data specifications that built upon national standards and other successful APCDs' requirements.**

Like many states, Kentucky has a "clean claims" law (KRS 304.2-110(1) and regulations (806 KAR 17:360) that requires carriers to pay most claims within 30 days of receipt. While some APCDs seek to collect member race, ethnicity and language data elements, these are not required for payment. The APCD should clarify with carriers what they are and are not able to provide. Carriers who submit to the Colorado APCD note that the data submission requirements recognize the types of data that are actually available in a carrier's systems. In addition, while data specifications should be uniform for all submitters, the standards for each submitter's level of completion may differ based on their organizational capacity.

3. **APCDs should provide strong "customer service" for data submitters during the implementation period and on an ongoing basis to facilitate problem solution.**

APCDs should establish customer service processes managed by informed individuals who can answer carriers' questions about data submission. Colorado and Rhode Island are good models for this, as both states have won high praise for being available to answer questions from their data submitters. Customer service is especially critical during the startup period for data collection; some states implement weekly office hours and other forums to answer questions and troubleshoot issues. States may find that these feedback loops are less necessary during the production period.

4. **Data quality should be checked and reported at every step in the process. This requires significant insight and experience with claims data.**

APCDs require a robust quality assurance (QA) cycle, with built-in checkpoints in the short term (at data intake) and long term (quarterly and annually). Short-term quality checks should be done before and after a carrier submits a data file. Checking incoming data files is typically an automated process and appears deceptively simple – for example, ensuring that each data field contains the correct values or number of letters. Data collection companies run several hundred checks. However, this initial process of checking the files requires expertise with claims data to define pertinent business rules, regardless of whether data intake is managed by an outside vendor or by in-house staff.

In addition, data checking at intake must occur within each file as well as across all files. An APCD data submission typically contains a member file, medical claims file, pharmacy file, and in some cases a provider file. The APCD should confirm that individual members are aligned across all files. Even when data submission rules are clearly defined, it is still possible for submitted files to contain claims records that are missing a corresponding member record. Therefore, thorough data checks are necessary to identify and address these discrepancies upon intake.

As the process continues beyond initial data intake, quality checks are needed on an ongoing basis. Even after files pass initial data checks, additional post ETL and production quality processes must examine consistency of the compiled information across all months in the reporting period for each data submitter. Examples of first line tests include month-by-month member counts, claims lines per

member per service month, per member per date of service month and per paid month, member and provider zip code distributions, and procedure and diagnosis frequencies.

Quarterly checks of overall membership and claims numbers are helpful in flagging any significant or unexplained increases or decreases that would skew analyses. Through such a quality check, the Massachusetts APCD was able to identify a carrier that had submitted multiple copies of the same enrollment file, thus skewing the results. Another long-term check is to compare the membership data in the APCD to those data reported by carriers to the insurance regulator in annual filings (as mandated by ACA law). Major discrepancies can be flagged and addressed with the data submitters. Building a thorough QA process to check the data on an ongoing basis will help the APCD team flag and address issues early.

State insurance departments can also assist APCDs in improving data quality. Two states cross checked APCD member totals against Rate Review submissions and identified discrepancies between the two data sources. State insurance department and APCD staff met privately with the data submitter's financial and reporting teams that are responsible for Rate Review filings and APCD submissions, respectively. The discussions spurred clarification of APCD data submission requirements and the data submitter's internal processes. The data submitters revised the APCD report specifications and resubmitted the files to provide data on previously omitted members.

Reporting to data submitters on whether their data has met the APCD's standards is a key feature of every state. Some states release publicly available data quality reports on carrier data submissions in an effort to improve transparency and data quality (an example of this is the MA APCD Data Profile Reports). It is important to note than in a voluntary submission model, carriers will have no incentive to meet the APCD's quality standards. Providing a quality data report would be useful but may not result in action taken by the carriers to remedy any deficiencies that are identified.

5. Voluntary data collection models may limit data quality.

The most effective data collection model among existing APCD states is one that is legislatively mandated. Mandatory data submission allows states to obtain standardized data on a regular, established basis. States can set clear expectations via rule for the submission process, including the required data specifications and a timeline for data submission with adequate test intervals. In addition, mandatory submission ensures that all carriers comply – thus creating a level playing field. National carriers (and those interviewed for the KyHDT) express a preference for mandatory data submission because a statutory mandate creates a clear legal path to providing the data under the HIPAA "public health" exemption. Carriers also note that a legal mandate helps justify the staff needed to handle the additional workload.

Voluntary data submission has proven to be a time consuming and, ultimately, unsuccessful model. Voluntary data use agreements can take at least one and a half years to negotiate and execute, and many never reach that point. Arkansas initially pursued a voluntary submission model for fourteen months without progress, before introducing a state mandate in 2015. Washington State has recently passed legislation to transition from a voluntary to mandatory model. Virginia's APCD is nominally voluntary; however, if the number of participating carriers falls below an established threshold, mandatory submission requirements become activated.

6. **The Data Warehouse architecture must reflect the proposed reporting needs.**

Once the ETL process has been completed and QA'd, APCDs usually run "value add" processes to improve speed and query efficiency. Processes run at this stage include:

- age bands
- grouping inpatient claims into a single stay and flagging associated physician claims
- geographic codes (county, census tract)
- condition categories
- primary care attribution
- carrier information (e.g., Medicaid, Commercial, Medicare)
- relative risk (illness burden) scores

Researchers and users access data through data extracts, data enclaves, and business intelligence (BI tool) layers. Selection of the particular BI tool must be evaluated against the designed reporting objectives. Some APCDs' reporting platforms are geared toward internal "non-public" reporting and analytics, whereas others are geared toward public facing reports. Different audiences and user groups may require more than one distribution model. For example, Colorado's APCD delivers non-PHI information through data extracts prepared by the APCD Administrator using a BI layer on top of a specially designed data warehouse. In contrast, Rhode Island's "power user" model gives each participating state agency access to a data mart to create agency-specific custom reports.

7. **Administrative simplification reduces the burden on carriers and increases operational complexity.**

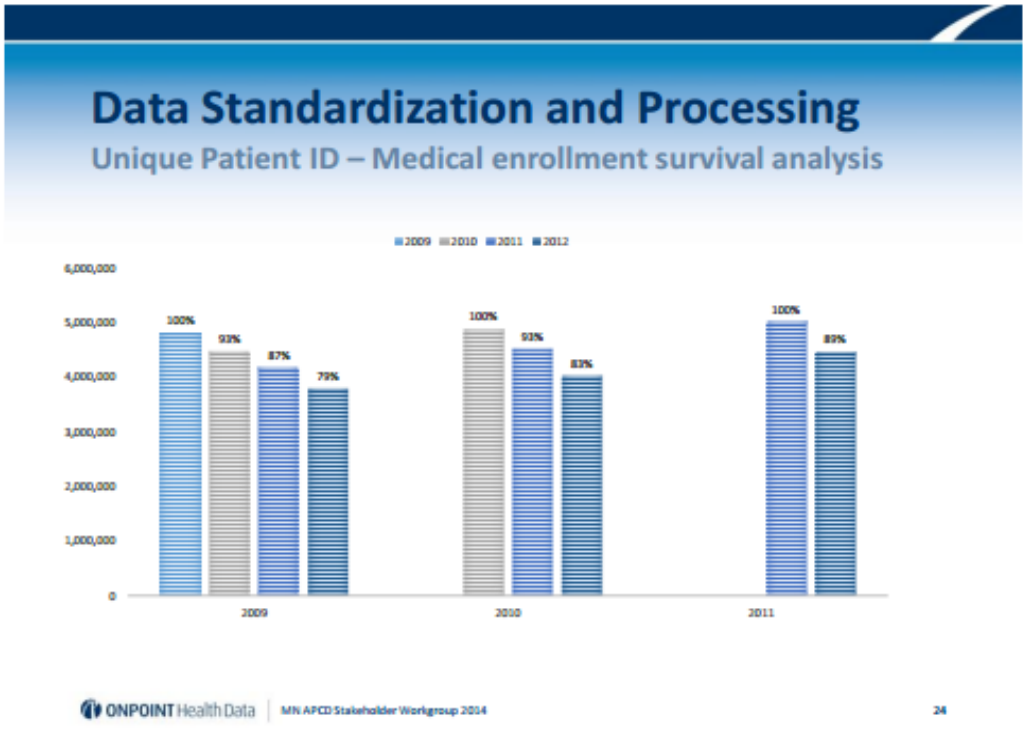
Administrative simplification refers to using one consolidated data feed from each carrier to meet the individual reporting requirements of multiple state agencies. While administrative simplification streamlines and standardizes carriers' reporting tasks, the APCD must in turn manage a complex array of data reports and timelines. State agencies will most likely have different formats and frequencies for their required reports; the APCD will therefore need to expand the types of data checks that are used. Data checking must be tailored to each participating agency's concerns and data requirements. Other new APCD tasks include building and maintaining an interagency governance structure to establish parameters on who may access which types of data and for what purposes.

8. **Strive to collect identified member data.**

A database with de-identified data will be unable to track individuals reliably for more than two to three years. Minnesota's APCD has stringent prohibitions on collecting identified data. Submitted data contains a hashed string of member name, date of birth and plan identifiers that cannot be parsed into component parts. This method usually provides consistent member information for a one- to two-year period. As shown in Figure 1, 21% of the members appearing in the 2009 member file were not matched in the 2012 file. Causes of "vanishing" members – besides moving out of state or changing a name – include changing carriers, employers, or even the product within the carrier.[1] Collecting name, date of birth and social security numbers (when available) supports more robust analysis than intake of de-identified data. Note that APCDs typically create strong protections around sharing data with direct identifiers and defaulting to the "minimum data necessary" standard when permitting access to any data extract.

[1] Note that the fall-off seems to be accelerating for the members in the 2010 and 2011 files. This graphic is included in a presentation made by Onpoint Health Data to the Minnesota APCD Advisory Committee on September 30, 2014.
http://www.health.state.mn.us/healthreform/allpayer/apcdonpointpresentation093014.pdf

Figure 1: Persistence of De-identified Member Information over Time in Minnesota



9. APCD data management models continue to evolve.

Although the APCD space is certainly not a mature market space, there are already several examples of states going out to rebid their APCD implementations. These rebids will result in a new vendor duplicating the initial work, raising some questions related to the best approach for building and hosting. At a minimum, initial vendor contracts need to be written to ensure that all data, both raw and processed data, are fully owned by the State and can be transitioned to a new vendor. Alternatively, evaluating other hosting models, such as internal hosting, may allow states to migrate more easily from one vendor to another as this industry matures. In addition, hosting data internally would allow states to engage multiple vendors for specific projects (though internally managed models do require a more robust internal technical team, which some states find challenging to hire and retain).

There are several different models for implementing APCD collection, management, and reporting. Some states (MA, KS) operate their APCDs in-house through a state agency. Other states (CO, OR, AR) contract with a single outside vendor to manage all data collection, analyses, and reporting. Still others (RI and, initially, UT) employ two or more vendors to handle separate stages (collection, management, and reporting). Minnesota, Vermont and New Hampshire delegate some services to a vendor and perform analysis in-house. Each model has its own strengths and weaknesses – most of which are related to cost, timeliness of APCD implementation, and flexibility to adapt the system to the state's evolving needs and requirements.

When developing an APCD, states must determine the best fit for their unique needs, resources, and vision. For example, Colorado's initial deliverable was a deadline-driven, public facing website that demonstrated that data could be successfully collected and appropriately reported. A single vendor offered streamlined development from intake to launch. As the database evolved, the Colorado APCD team began to build its own reports from a BI tool provided by the data manager. Rhode Island, seeking to increase access to data and analytics in its human services and health agencies, is leveraging a web-based BI tool into its vendor-managed APCD data warehouse as a test of a distributed analytic capacity among participating state agencies. States need to balance the advantages of an integrated single contractor against the depth and breadth of expertise available through a multi-vendor array.

Each of the current APCD data vendors offers different strengths. Some have demonstrated deep expertise in data collection and file production while others offer more experience with analytic services. Key capabilities needed vary with the project's scope and may include:

- Knowledge of health care claims data processing
- Familiarity with the state's health insurance market
- Insight into the uses of health data to drive system change
- Experience with aggregating multi-payer claims data
- Collaborative approach to developing reports and analysis
- Tools and products that do not limit potential reports
- For public reports and websites, health literacy expertise

10. **Inventory needs and provide as much detail as possible about desired products and services before hiring a data vendor or building a data warehouse.**

States benefit from developing a clear and detailed Request for Proposals (RFP) for an APCD data vendor. The RFP should be fully transparent in the expected scope of services and desired deliverables to avoid costly change orders and delays. States should overstate their needs and be as detailed as possible to ensure that vendors' RFP responses clearly convey their capabilities for meeting the state's needs. In addition, the RFP should request information on how respondents would provide training for the state's APCD users on any tools or reports, as this is a critical component of APCD implementation. The RFP should also include a model contract that lays out a clear set of end-of-term expectations for intellectual property, transition to a new data warehouse, and original and processed files and documentation.