Best practices for launching and operating an all-payor claims database

Al Prysunka

INTRODUCTION

The high and sustained growth rate of healthcare in the United States. over the past two decades has created significant financial pressures for both government payors (Medicare/Medicaid/VA) and private sector employers who offer health insurance to their employees. Additionally, while per capita healthcare expenditures have increased at an alarming rate, many believe that the overall quality of healthcare in the United States is not commensurate with the expenditures and that significant improvements in health outcomes for the general populace have not been realized.

In order to address the cost and quality issues in the U.S. healthcare system, both government policy makers and employers have found it critical that a comprehensive and timely source of data be available to better define the problems and to set forth proposed solutions. The data set that has emerged to meet those requirements is the all-payor claims database (APCD). This paper will provide an overview of the structure and key considerations for planning, launching, and operating an APCD, and will list a number of potential uses for the data, once collected.

DEFINITION OF APCDs

APCDs are large-scale databases that systematically collect healthcare claim data from the existing transaction systems created to pay healthcare claims. The data can be derived from a variety of payor sources, including commercial (insurance carriers, third-party administrators, pharmacy benefits managers, dental benefits administrators) and government (Medicare, Medicaid, Federal Employees Health Benefits Program [FEHBP], Tricare). Because the data sources are so comprehensive, APCDs contain claim records from most healthcare providers (facility and practitioner) offering services to defined members.

The information typically collected in an APCD includes member demographics, provider demographics, clinical, financial, and utilization data. Most APCD systems typically include eligibility data and medical, pharmacy, and dental claim data. Denied claims are not typically captured due to the difficulty of reconciling resubmitted or replacement claims with the original. Because claims are often not created for services provided to the uninsured, data on this segment of the population are usually unavailable in APCDs.

APCDs can exist as a statewide, comprehensive database managed by an agency of state government or its designee. These databases are established through legislative action (with the promulgation of rules further defining the requirements) and generally cover most of the commercially insured population of a state and the larger governmental payors. The defined uses of the data can be broad in scope and the data and/or reports are usually accessible to the public.

APCDs can also be created at a regional or sub-state level, usually by nonprofit organizations such as healthcare-related employer business groups or community coalition organizations. Because of HIPAA requirements and legal issues with obtaining Medicare data for comprehensive uses (Medicare data are available for specific, defined CMS initiatives), these sub-state APCDs usually contain data derived from, or related to, members of their organizations and have more narrowly defined uses. The data are primarily released to subscribers and members, with some reports being made available to the public at large.

KEY ELEMENTS OF AN APCD

Governance

As mentioned previously, statewide APCDs are usually established through legislative action with corresponding administrative rules that define the data categories to be submitted, file and code specifications, frequency of submission, and data release policies. While most states have legislation mandating commercial healthcare payors to submit claim data (with penalty provisions for failure to submit), some states have crafted legislation creating a voluntary submission process, which remains voluntary as long as the payors are responsive to the needs of the state. A mandated approach increases the probability that the commercial part of the database will be more uniform and complete. However, because federal law supersedes state law, the inclusion of the government data (e.g., Medicare and Medicaid) is contingent upon states entering into formal agreements with the federal agencies responsible for the data. These agreements are written in accordance with federal law, regulations, and policies pertaining to the release of the data to third parties (including states) for various uses.

Sub-state APCDs are created voluntarily and function under the managing organization's charter and through data use agreements of the participating organizations (payors, providers, and employers). Sub-state initiatives may be more flexible in their operations and provide important information to their participants, but are usually limited in scope and are dependent upon the continued cooperation of the participants. To date, CMS has been reluctant to provide

Medicare data directly to sub-state entities for inclusion into an APCD for research (multiple-use) purposes. CMS has also been reluctant to allow states to provide Medicaid data for purposes not specifically related to the improvement of the program. However, sub-state entities can receive limited Medicare data sets to incorporate with their commercial data through specific CMS programs such as the Multi-Payer Advanced Primary Care Practice Demonstration and Performance Measurement pilots, and can receive Medicaid data under an agreement with the state government entity managing the Medicaid program.

Whether establishing a statewide or sub-state APCD, it is critical to obtain input from all stakeholders during both the development and operational stages. This can be accomplished through the establishment of formal boards or committees, with specified representation, through the creation of ad hoc committees with specific subject matter expertise, or a combination of both. Because of their technical and administrative complexity, APCDs cannot be built and operated successfully without open lines of communications among all stakeholders.

Funding

Funding the operational costs of a statewide APCD can be accomplished by using a number of different approaches. The most common sources of funding are: a general fund appropriation; dedicated revenue through assessments on carriers, third-party administrators (TPAs), and/or healthcare providers; and revenue derived from the sale of data. States have also used federal funding to finance the bulk of the costs or to provide supplemental dollars, includin: Medicaid 50% federal match of administrative dollar, and Exchange Establishment Cooperative Agreement grants available to help establish health insurance exchanges. Additionally, some states have pursued grants from foundations and trusts to fund particular aspects of APCDs.

Funding the operational costs of a sub-state APCD is usually accomplished by directly assessing the member organizations on a prorated basis. These dollars can be supplemented through grants from foundations, trusts, an, in some cases, contractual dollars from state governments. Federal grants are available for incorporating a limited Medicare data set into a sub-state APCD.

Data sources

As mentioned previously, data for APCDs can be derived from a variety of payor sources, including commercial (insurance carriers, TPAs, pharmacy benefit managers [PBMs], dental benefit administrators) and government (Medicare, Medicaid, FEHBP, Tricare). All APCDs contain data from insurance carriers. However, because some states do not license TPAs and/or PBMs, it is difficult to determine which companies are paying claims in a given state. Consequently, the data from these entities are not captured in the database. In addition, thus far it has been difficult to obtain data from the FEHB, Tricare, and the VA. This means there is almost no information about federal employees or active and retired military personnel. Only one state has acquired FEHBP data, and no states have acquired any Tricare or VA claim data.

States have also deliberately limited the number of commercial healthcare payors submitting data, which is due to cost considerations. The number of commercial payors, and the percentage of the state market held by each, differs significantly from one state to another. A state may have hundreds of licensed payors, but the vast majority of the market (95% or more) may be covered by 50 or fewer. To collect data from the remaining payors would not be cost-effective. Consequently, limits have been developed, usually through rules, to exclude payors from submitting data. Limits can include a minimum number of members per month and/or a minimum dollar amount of premiums written or claims paid annually. This information is usually captured through a registration process or is collected separately by a state insurance department.

Exclusions/exemptions

In addition to excluding payors from submitting any claim data, exclusions are provided for specific types of policies or products offered by a healthcare payor. These exclusions are created because the policies or products are missing key components of an APCD. The most common deficiencies are lack of an eligibility file and the inability to accurately track or link claims for an individual over time. Examples of these types of policies are workers' compensation, long-term care, specific diseases, and student coverage.

Most statewide APCDs only allow data related to adjudicated claims to be submitted and prohibit the submission of denied claims. This is because denied claims are resubmitted by the providers until the problems are resolved and the claim is paid. Every resubmitted claim has a different claim ID and it is very difficult to find and purge the earlier submitted versions from the system to avoid duplicate service counts.

Because there are no claims created for the uninsured, a significant percentage of the population is omitted from an APCD. However, one state is collecting pseudo-claims, which are created by a TPA for a hospital system for some of its uninsured population.

Data collected

All APCDs contain eligibility files and medical, pharmacy, and/or dental claim files. APCDs also have unique member and provider ID tables created from the data submitted. The unique member tables can be identifiable but, because of the sensitive nature of these data and privacy concerns by the public, the data is usually stored and utilized in an encrypted form. The provider tables contain identifiable data. However, accurately identifying providers across payors in the data has been difficult, which is due to the indiscriminate substitution by the payors of the rendering provider with the billing provider. This has also impacted the ability to assign individual practitioners to group practices or hospitals and other facilities. Provider identification has also been complicated by requiring physician assistants and nurse practitioners to use the national provider identifier (NPI) of the supervising physician, resulting in ambiguous attestation.

With respect to the individual data elements collected in an APCD, those shown in Figure 1 are usually required.

FIGURE 1: APCD INDIVIDUAL DATA ELEMENTS

Encrypted (or unencrypted) subscriber/member names, SSNs, and plan assigned contract number	Facility type (hospital, SNF, surgical center, etc.)
Type of product (HMO, POS, indemnity, etc.)	Service and billing provider information (names, NPIs, tax IDs, specialty codes, location/affiliation)
Type of contract (single person, family, etc.)	Prescribing physician identification
Patient demographics (date of birth, gender, residence)	Bill type
Diagnosis codes (including E-codes)	Revenue codes
Procedure codes (ICD, CPT, HCPC, CDT)	Plan payments
NDC code/generic drug indicator	Member payment responsibility (copay, coinsurance, deductible)
Paid date/service dates	

In an effort to reduce data submission and processing costs, and to standardize the data sets across state boundaries, two national standards organizations have produced reporting guides or standards for APCDs.

On October 25, 2011, the National Council for Prescription Drug Programs (NCPDP) released the Uniform Healthcare Payer Data Standard, which creates administrative efficiencies and supports the reporting requirements for pharmacy claim data submissions to states or their designees. It also established criteria to be used for all entities sharing historical pharmacy-related healthcare data.

In addition, the Accredited Standards Committee (ASC) X12 has also approved Institutional, Professional, and Dental Post-Adjudicated Claims Data Reporting (PACDR) guides for publication in October 2012. These guides were developed by a Special Appointed Committee that will continue to meet to develop standards and implementation guides for the eligibility file component of an APCD.

Although it is not mandatory that the reporting guides be followed, it would be desirable from a cost and operational perspective for any state or sub-state entity operating or establishing an APCD to utilize the reporting guides.

Because of the content and limitations of healthcare claims and remittances, not all data desired by those establishing APCDs can be included within the data files submitted by payors. Data elements typically excluded from an APCD are shown in Figure 2.

FIGURE 2: DATA EXCLUDED FROM APCDS

Premium information	Referrals
Capitation fees	Provider networks
Administrative fees	Clinical results from lab work, imaging, etc.
Back-end settlement amounts	

BEST PRACTICES FOR APCDs

Technical requirements

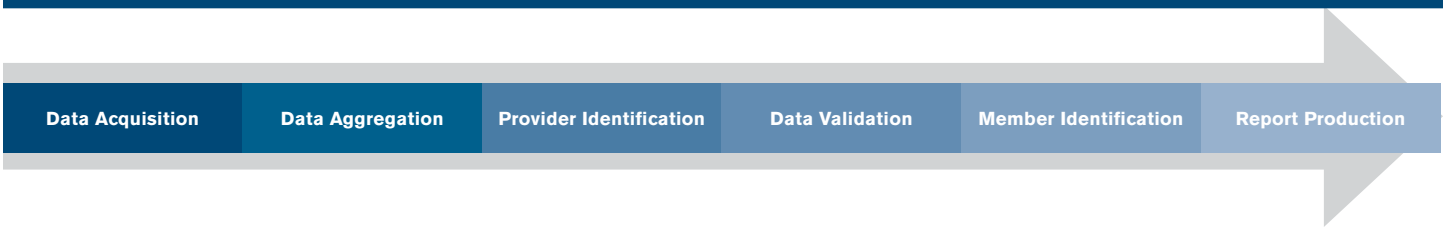
Because of the large number of data submitters, data files, elements, codes, and also the large size of APCDs, the system design (Figure 3) for receiving, loading, processing, storing, and retrieving claim data should be developed with considerable thought.

A registration process for all data submitters needs to be constructed in order to ensure that all of the required claim data files are being submitted. To maintain a high level of data integrity and quality, it is necessary to create load and quality edits that are applied to each element as the data files are submitted. Because of the large size and number of files submitted, the edits are usually designed in a manner that rejects the entire file if an individual element fails to meet the specified threshold.

Consideration must also be given to the design of the database as the data is loaded, stored, and retrieved. APCDs are usually built as relational databases, utilizing cubes and/or segments to better access the data when needed. The database design must also take into consideration the need to have multiple years of data stored and the ability to alter and/or replace adjudicated claims over time.

Because of the size of an APCD (many millions of records per year that accumulate over time), the system must be designed with multiple terabytes of data storage and servers with significant processing power. Whether the hardware resides on-site, at a vendor, remotely through the cloud, or some combination of the three will depend on the needs of the individual state or sub-state entity.

FIGURE 3: CLAIM DATA PROCESSING FLOW DIAGRAM



If the data are to be used (as files or processed tables) by entities in addition to the state or sub-state organization collecting the data, thought must be given to how the data are to be accessed or displayed. Data can be accessed through a website, secure file transfer protocol, or DVDs. Tables, dashboards, and other products created from the data can be viewed on public websites or can be created by outside users via a secure portal that has limited access through subscriptions or registrations.

Linking of APCDs to other databases

As APCDs are being created, the next step in the progression of obtaining more comprehensive, timely, and accurate healthcare data is to link the APCDs with other databases. The three databases most commonly discussed include two existing databases (hospital administrative and vital statistics) and data from the newly created health information exchanges (HIEs). It is critically important that the APCD be constructed with data elements that allow linkage to the other databases, and that identical elements exist in the other three databases. These elements include: member/patient names, dates of birth, Social Security numbers (SSNs), patient account/control numbers, dates of service, and provider identification (NPIs, names, locations).

Although hospital claims exist in APCDs, the data content is not as comprehensive (fewer diagnosis and procedure codes and groupers) and hospital databases cover all patients (including the uninsured). Conversely, hospital databases contain charges but no payment information. Where hospital and claim databases coexist and both databases are available to the public, consideration must be given to releasing charge and payment data simultaneously at a detailed level, which could result in the exposure of contractual discount rates.

Linking APCD data with vital statistic data (e.g., cancer registry data) has great potential to improve the understanding of disease prevalence rates in a particular service area, and to analyze the quality of treatment received by those patients in the same areas.

HIEs have the potential to enhance existing APCDs with clinical information for quality and outcomes reporting. Because of their content and purpose, HIEs and APCDs will be distinctly separate initiatives as they are developed. If both are integrated, data will exist for comparative effective research, for population health applications, and to improve risk adjustment, clinical studies, and outcomes research.

One large impediment can exist to make this linkage extremely difficult, if not impossible. HIEs can be designed in two basic models: distributed and consolidated. In a distributed model, the data resides with each provider and is pulled across the system only when needed. In a consolidated model the data moves from the individual providers to a central data repository (which can easily be converted into a database) and then to the users.

In order for APCD data to be linked to a distributed HIE, the linkage would need to occur with every individual provider, which would make it technically difficult and cost-prohibitive.

Uses of APCDs

APCDs provide transparency across the entire spectrum of healthcare payors and providers. Such transparency enables a wide variety of stakeholders to access information that provides insight into how and where healthcare dollars are spent. APCDs have the potential to improve provider quality of care, to allow payors to create incentive systems to reward the delivery of high value and efficient care, and to provide consumers with information to make rational healthcare choices based on cost and quality.

Thus far, data from APCDs have been used for the following purposes:

- Providing price information to consumers for specific services available through public websites
- Establishing cost and utilization rates by geographic regions
- Producing quality evaluations of individual healthcare providers (including the development of quality metrics)
- Evaluating individual members' healthcare across payors (including the analysis of the Medicare/Medicaid dual-eligible population) and across all provider categories (including the use of episodic treatment groups)
- Evaluating disease prevalence across a specific population
- Conducting payment reform models (accountable care organizations, patient-centered medical homes)
- Analyzing payor competiveness within the commercial insurance market
- Supporting federal initiatives (including the health insurance exchange reinsurance and risk adjustment processes, the CMS Multi-Payer Advanced Primary Care Practice Demonstration pilots, and the CMS Performance Measurement pilots)
- Determining Medicaid fraud and investigating provider anti-trust activities

THE FUTURE OF APCDs

The United States cannot continue to provide healthcare in the same manner it has been—it is simply not affordable. The need for timely and accurate data to provide a basis for healthcare reform will continue. Hospital databases are too limited in scope and HIEs, although extremely valuable in collecting clinical data, have not yet been fully developed, which is due to legal and cost considerations, and they do not provide payment information. Even with their deficiencies, APCDs are still the only viable option to offer a comprehensive, cost-effective look at the healthcare delivery system. Because of this fact, APCDs will continue to be created, both at the state and sub-state level, for the foreseeable future.

Al Prysunka is a senior consultant in Milliman's Seattle office and director of the MedInsight all payor claims database products. Contact him at al.prysunka@milliman.com.