

Final Contract Report

Inventory and Prioritization of Measures To Support the Growing Effort in Transparency Using All-Payer Claims Databases

Prepared for:

Agency for Healthcare Research and Quality
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. HHS2902001200003I, Task Order 5

Prepared by:

Stanford University (Prime Contractor)

Sheryl M. Davies, M.A.
Kathryn McDonald, M.M.

Truven Health Analytics

Erica Danielson, Ph.D.
Manju Gokhale, M.A.
Stella Chang, M.P.H.

University of California, San Francisco

Naomi S. Bardach, M.D., M.A.S.
Grace A. Lin, M.D., M.A.S.
Mitzi Dean, M.S., M.H.A.
R. Adams Dudley, M.D., M.B.A.

APCD Council

Denise Love
Ashley Peters
Josephine Porter
Emily Sullivan

Stanford University

Center for Primary Care & Outcomes Research, School of Medicine
Center for Health Policy, Freeman Spogli Institute for International Studies

AHRQ Publication No. 17-0022-1-EF

March 2017



The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

Disclaimer of Conflict of Interest

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Funding Statement

This project was funded under contract number HHS A2902001200003I from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services. The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the U.S. Department of Health and Human Services.

Public Domain Notice

This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated.

Suggested citation: Davies SM., McDonald K, Danielson E, et al. Inventory and prioritization of measures to support the growing effort in transparency using all-payer claims databases. Prepared under Contract No. HHS A2902001200003I, Task Order 5. AHRQ Publication No. 17-0022-1-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2017.

Table of Contents

Introduction.....	1
Overview of All-Payer Claims Databases	1
Uses of All-Payer Claims Databases	1
Purpose of Project.....	2
Reviewing the Landscape of All-Payer Claims Databases	3
Technical Expert Panel.....	3
Environmental Scan.....	6
Measure Inventory	11
Data Evaluation of All-Payer Claims Databases.....	12
Methods.....	13
Results of Data Evaluation	13
Discussion	21
Limitations of APCD Data for Measurement	22
Ongoing Efforts To Improve APCD Data	24
Conclusion and Future Directions	25
References.....	27
Appendix A: Technical Expert Panel and Learning Network Members.....	29
Appendix B: Prioritization of Conditions for Measure Inventory	30

Introduction

Overview of All-Payer Claims Databases

For the past three decades, measures of health care quality have been embraced as tools for quality improvement, transparency, and accountability. As health care costs continue to rise, more stakeholders are calling for such measures to be used as part of value assessments, to understand how high health care expenditures, high out-of-pocket costs, and high local and Federal spending on health care can be shifted toward better care at better prices.

Price transparency, or readily available access to information on variations in cost across health care providers, alongside access to quality metrics, may help shape innovative policies, improve health care efficiency, and promote value in the health care system. Comprehensive datasets that support the development of quality and cost measures across large populations are well suited to reach these goals.

All-payer claims databases (APCDs) incorporate utilization and cost across the health care system and may fill critical data needs for State agencies and other stakeholders. The hope is that these databases can serve to inform policy and provide transparency for decisionmakers, including consumers, purchasers, and policymakers. APCDs are large datasets that aggregate medical, facility, pharmacy, and sometimes dental claims, as well as eligibility and provider data from private and public payers.¹

Currently, 18 States have legislation mandating the creation and use of APCDs or are actively establishing APCDs. More than 30 States are maintaining or developing an APCD, or have a strong interest in developing an APCD. Seven States have public reporting Web sites with cost and quality information either wholly or in part coming from APCD data.¹

Uses of All-Payer Claims Databases

APCDs represent a new approach to providing information about care. Before the advent of APCDs, datasets tended to be limited to certain groups (e.g., Medicare and Medicaid claims, which omit a large portion of the overall population) or to particular components of care (e.g., hospital discharge abstract databases, which cover only inpatient care) and posed challenges making it difficult or expensive to follow patients longitudinally.

The promise of APCDs is that they allow creation of a more comprehensive picture of care than is otherwise available in most States.² By collecting data from all payers, State APCDs capture encounters for all but a small minority of patients (e.g., patients who are uninsured or are covered by a Federal health care plan such as Veterans Affairs benefits) and across settings.

This expanded database has several advantages, including that patients in a well-implemented APCD can be followed over time and across settings, to capture full episodes of care and account for variations in the type of care received. APCDs are not limited by turnover in patients among providers or payers because the records are captured for each patient regardless of provider or payer. This unique aspect of an APCD can facilitate measures of continuity of care, coordination of care, and other traditionally difficult constructs to measure.³ Compared with single-payer

databases, APCDs may have larger sample sizes, potentially facilitating measures for rarer events and among smaller entities (e.g., individual providers, small areas).

APCDs have a variety of other potential uses as tools for improving quality of care and population health. As mentioned at the outset, they create a more comprehensive picture of outpatient and pharmacy care than is otherwise available in most States as they include a large percentage of insured individuals, which allows APCDs to reflect the health of the population.²

APCDs also facilitate price transparency and highlight significant price variation in the system. Such transparency may allow purchasers to negotiate with providers more effectively. In addition, it can help providers assess their own quality and value in shared risk and accountability payment models and allows consumers to weigh value in health care decisions as they assume greater financial responsibility.⁴

Local data aggregated into APCDs can be used to understand local market functioning and assess whether spending variations reflect pricing, utilization, or both. APCDs also can provide data to help States develop strategic plans for public health legislation or to determine the impact of policy changes at the State level.⁵ In addition, APCDs can support research that may be of interest to State policymakers, such as comparative effectiveness studies or the development and evaluation of targeted interventions to improve chronic disease care.

The comprehensive nature of APCD data allows:

- Estimations of disease prevalence across a population,
- Identification of patterns of utilization and potential areas for targeted interventions, and
- Planning and evaluation of health reform programs and legislation on cost, quality, and access to care.

While APCDs offer many advantages over other databases, they do have known limitations. These include lack of data for uninsured patients, questions about the feasibility of gathering and maintaining datasets, variation in data quality among submitters, and lack of clinical detail (e.g., laboratory values, biometric details). It is unknown how these limitations affect the ability to use APCD data for quality and value measures locally, within States, and nationally.

Purpose of Project

The Agency for Healthcare Research and Quality (AHRQ) funded the APCD project, which focused on three areas:

- Review of the overall landscape of APCDs;
- Identification of measures of cost, utilization, and quality that can be defined using APCDs; and
- Evaluation of current APCD data for measurement use.

In addition, the project team convened a Technical Expert Panel (TEP) to provide input on the overall direction of the project and provide check-in opportunities on key project milestones. The TEP provided input for the environmental scan and literature review that was conducted for the

measure inventory. The TEP also provided critical feedback on priorities for the measure inventory.

The project focused on measures that address aspects of the health care system not covered well by other datasets, such as ambulatory care, pharmacy, and cost. Multiple use cases were considered for APCD-based measures:

- Supporting States in achieving overall high-value care for their population.
- Assisting physicians and physician groups in understanding quality and cost in managing their own patient populations.
- Supporting consumer choice when choosing providers or health plans, although currently this application is limited by the lag in price and quality information, and in some cases lack of relevant measures to support the decision.

After internal deliberations and TEP consultations, the team agreed to focus on measures that supported population health management.

This APCD project resulted in three specific work products:

1. **A report on the current APCD measurement landscape:** The goal of the environmental scan was to assess potential use cases for APCD measures and to understand and summarize the current evidence for and limitations of APCD measures. The environmental scan informed both the measure inventory and data analytics.
2. **An inventory of measures that could be constructed using APCD data:** The overall goal of the measure inventory was to provide a useful and usable inventory of measures that could be derived from APCDs, to provide a framework and basic measure specifications for assessing existing measures for use with APCDs, and to inform future measure assessment for other APCD users and use cases.
3. **A preliminary analytic assessment of three existing APCDs:** The goal of the data assessment was to acquire three APCDs and assess data availability, documentation, and data completeness. The analysis focused on aspects that were relevant to measures and provided a preliminary assessment of the completeness and face validity of critical data elements.

Reviewing the Landscape of All-Payer Claims Databases

Technical Expert Panel

Panel composition and role. A Technical Expert Panel (TEP) was formed based on suggestions from the project team and the AHRQ project officer. The TEP consisted of State APCD representatives, researchers, and consumer advocates, as well as a learning network (Appendix A lists TEP and Learning Network members). TEP members provided input on overall project direction and check-in on key project milestones during three focused TEP calls and occasional email-based discussions.

Goals for the first TEP call were to garner feedback on the approach to the APCD measure inventory. TEP feedback was desired on the framework for organizing and prioritizing measures.

The second TEP call used a thought experiment to help increase the usability and relevance of the measure inventory. The third TEP call focused on high-priority clinical areas to focus a more intensive evaluation of select potential measures (“the deep dive”).

Over the three TEP calls and emailed feedback, the TEP provided key information on the APCD uses and users, organization and prioritization of measures, ideal application of APCDs to improve population health, and limitations of APCDs. They also ranked APCD use cases from various condition-specific areas.

APCD measure audiences and potential uses. TEP panelists noted that APCD measures have a variety of potential audiences, ranging from policymakers to consumers. Potential audiences included State health and data agencies, policymakers, payers, providers, third-party service providers, data or business analytic companies, consumers, and researchers.

Panelists maintained that policymakers were the primary target audience for initial reporting of measures using APCD data. These policymakers can affect further development of APCDs, as they need population-based measures that use readily available data, and are less affected by limitations such as data delays, which can severely limit the use of measures in consumer choice and negotiation.

Panelists noted that among the uses discussed, including choice, negotiation, accountability, and policymaking, one of the biggest potential targets for APCD measures is the market at large. Price transparency can be a powerful tool to help even out variations and constrain health care costs.

To further prioritize potential measures for inclusion in the measure inventory, panelists discussed the usefulness of the National Quality Forum (NQF) list of 20 high-priority Medicare conditions as one potential resource. Panelists noted that it may not adequately capture the interests of diverse audiences and recommended that consideration be given to three distinct payer audiences: Medicare, Medicaid, and payers. Consumers, for instance, would need measures that are “shoppable,” meaning they cover conditions for which consumers have time to investigate the value of provider options.

The consumer use case was repeatedly discussed because consumer audiences have increasingly become a focus in reporting and use of APCDs. Panelists felt that while it would be useful to provide consumers with information on out-of-pocket costs, this use case presented certain challenges. A weakness of APCDs is the lag between claims and what is currently happening in the market. Given this delay, using the APCD for current pricing may not be effective or appropriate. Furthermore, informing the consumer was noted to be difficult and it would be hard to reach large numbers of consumers to change their behaviors in meaningful ways. Ultimately, an informed decision was made to not integrate consumer choice into the measure inventory.

Finally, usefulness and use cases are affected by data consistency and the ability to apply the measures across different APCDs and for across-State comparison. An illustration of this issue is that not all APCDs use the same clinical groupers (level of major diagnostic categories, diagnosis-related groups [DRGs]), and some do not use groupers at all. Thus, measures based on these groupings cannot be applied across APCDs.

Framework for organizing measures. Panelists expressed support for a framework presented by the project team, which organized measures by three functions of measures: choice, negotiation, and accountability. However, they noted that this might not be the most effective way to organize measures within the inventory. Panelists expressed the need for a framework that was flexible and could encompass the many possible current and future uses of APCD data.

Ideas for improving the framework included:

- Adding uses or functions such as feedback to providers and transparency.
- Considering which audience or stakeholder a use case applies to.
- Considering who the individual users are and their unique needs.
- Integrating use (such as the Choice, Negotiation, Accountability framework) and audience.

Ideal application of the measure inventory. To guide the presentation of the measure inventory to maximize its usefulness for the target audience of States, TEP members were asked to engage in the following thought experiment: “Imagine you are a state developing and/or maintaining an APCD. What would be the ideal application of that APCD to improve population health?” In response, TEP members provided the following feedback:

- TEP members noted the power of an APCD dataset lies in its ability to look at a broad cross-section of patients from multiple different payer types for any given provider. It was suggested that there is great interest in using such data to spot variation (e.g., in utilization, spending, rates of preventive screenings) and determining why that variation exists.
- Panel members expressed interest in looking at children across commercial versus Medicaid and the Children’s Health Insurance Program to analyze access and utilization rates to study disparities.
- TEP members found it valuable to make data available to researchers and to have a dataset that allows people to investigate a broad range of topics.
- Topics of great interest included high-cost populations and the opioid addiction problem.
- Finally, panel members suggested that it might be more manageable to classify different stakeholder groups, develop an understanding of how each contributes to population health, and determine what information might allow each to make more meaningful contributions to population health.

High-priority clinical areas for measurement. To assist with selecting measures that would be assessed in depth, TEP members discussed and ranked the importance of use cases from a predetermined set of clinical priority topics: Diabetes, Mental Health and Substance Abuse, Childbirth and Reproductive Health, Imaging, and Medications. These topics met the criteria of being common clinical concerns and having multiple identified potential measures.

- **Diabetes:** Some members ranked this condition as high priority due to the large proportion of health care spending and it being a broad public health problem. Diabetes also offers an opportunity for improvement by better understanding differences in patterns of care across geographic regions.

- **Mental Health and Substance Abuse:** Members had widely differing opinions on Mental Health and Substance Abuse. One member ranked it as high priority because the topic was of interest in his/her State, while others ranked it as middle or lower priority. There was also uncertainty among the panel about the collection of mental health and substance abuse data, especially, privacy laws and concerns or noncoverage of care. Although some saw this condition as a high priority, others raised concerns that measures identified in the measure inventory were not sufficient to understand the cost/quality paradigm, specifically due to the limited number of quality measures. Also, utilization and cost measures in the absence of quality measures could be misleading and thereby discourage people from seeking services altogether in areas with limited choice or alternatives.
- **Childbirth and Reproductive Health:** Some members felt this area was high priority, because from the consumer perspective, there is time to plan; from the provider perspective, there is an opportunity to inform and educate. It was noted that the frequency of the ongoing prenatal care measure would be a challenge and APCDs may not be able to support this measure, but the area crosses payers, thereby providing consumers the ability to take action.
- **Imaging:** This use case ranked low among TEP members. They noted a limited breadth of measures, along with issues such as sizable price variation in imaging; and different technologies used for the same procedure, which leads to unproductive conversations and disagreement on costs.
- **Medications:** One of the advantages of an APCD is the ability to highlight differences in amounts paid for common medications. There are opportunities to focus the measures around medication to inform public policy that could result in better care and lower costs.

More TEP members prioritized diabetes than other categories, stating that the topic has the prospect of improving care dramatically for a significant portion of the population. In addition to diabetes, members suggested expanding the topic to focus on high-cost, manageable, chronic conditions in general (e.g., asthma, mental health, substance abuse). Mental health was the second most commonly prioritized topic area; TEP members hoped that the focus would promote development of better measures.

Environmental Scan

Objective. The objective of the literature review and environmental scan was to map an approach to creating an inventory of measures of quality, cost, and utilization of care across settings for potential use with APCDs, noting gaps or current barriers to APCD measurement. The literature review and environmental scan provided a foundation for the work, describing the breadth of available measures and generating a framework for choosing measures and organizing them in the final inventory.

A large number of measures are based on administrative health data, in particular, hospital discharge abstract databases. For the purpose of the environmental scan, the focus was on measures that leveraged the unique aspects of APCD data. These included longitudinal data from multiple sources that allow patients to be tracked across time and settings, pharmacy data, and data on dollar amounts paid by insurers and patients. These key characteristics of APCDs enable measurement not possible with hospital discharge claims data alone, in particular, measures of

ambulatory care (including measures that require data from multiple settings), episode-of-care measures, and cost measures.

Guiding questions. Several guiding questions were used to focus the search strategy and data collection efforts. The guiding questions resonated with the objectives of the report:

1. What measures or outcomes (quality, utilization, safety, price, etc.) that leverage the unique data in APCDs have been reported in the scholarly literature or in online public reports using APCD data?
2. What measures or outcomes have been proposed for use with APCD data or claims data that are episode based or longitudinal in nature?
3. What important measure gaps have been noted in relation to transparency initiatives? Have APCD-specific measure concepts been proposed to fill these gaps (even if no fully specified measures yet exist)?
4. What potential barriers to using and reporting measures with APCD data have been identified in the peer-reviewed or grey literature, including issues around availability and access to data elements? What strategies for overcoming these barriers have been proposed in the literature?
5. What are some of the methodological considerations and issues pertaining to using APCD data for measurement that have been discussed in the peer-reviewed or grey literature?

Summary of key findings from literature review and environmental scan. The peer-reviewed literature search yielded two basic types of articles: overview papers presenting concepts related to measurement in the areas of focus and those presenting particular measures used in addressing the authors' specific research question. The team screened a total of 189 articles; of those, 98 articles were included in a full review, of which 17 were overview articles and 81 were research articles.

Overview articles presented concepts related to measurement using APCDs or other large claims databases. The research articles focused on measurement of cost, quality, or utilization using APCDs, multipayer claims databases, or other claims databases (Medicare or Medicaid). The measures found most often focused on a specific research question or discussed the application of specific measures for public reporting or price transparency initiatives.

The environmental scan yielded information on existing APCDs and the potential of using APCDs to measure quality, cost, and utilization of care across populations and settings. Of the 236 different sources of information identified using the initial search criteria, 127 sources were included that met the inclusion criteria for this report. Sources found in the environmental scan came from AHRQ expert materials, the APCD Council, reports, task force papers, policy briefs, trade papers, business journals, white papers, books, APCD public reporting Web sites, measure inventories such as the National Quality Forum (NQF), and other sources.

National and State-specific general resources most often described the basis for the national trend to develop APCDs, as well as State-specific issues related to building or implementing APCDs. State resources, in particular, focused on building the case for and the barriers to establishing APCDs and using them for measurement. Many States with active APCDs have also issued reports of statewide quality, utilization, and cost.

Finally, resources related to price transparency were also included, as this is a major application of APCDs. The resources included high-level summaries of current efforts and barriers from governmental and nongovernmental sources, as well as State-level price transparency reports.

Major organizations and key contributors to APCDs and measurement. Several national organizations have developed expertise with APCDs and measurement using APCDs and have created resources and provided support for States creating APCDs. Such organizations include **AHRQ**, which has supported several initiatives, including:

- The United States Health Information Knowledge (USHIK) database (available at <https://ushik.ahrq.gov/mdr/portals>), a repository for State APCD file submission specifications and data elements;
- The Community Quality Collaboratives program (information available at <https://www.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/localnetworks/index.html>), where States shared experiences and best practices in quality and efficiency measurement, public reporting, and transparency; and
- The National Quality Measures Clearinghouse (NQMC) (available at <https://qualitymeasures.ahrq.gov/>), an online, searchable inventory of evidence-based measures and measure sets.

Nongovernmental organizations have also played a key role in advancing the science and implementation of APCDs:

- The **APCD Council** and the **National Association of Health Data Organizations** have taken the lead in supporting and documenting current State efforts and legislative work around APCDs. Efforts include working to harmonize data collection and release across States and providing technical and policy support to States that have or are developing APCDs.
- **Catalyst for Payment Reform** is a nonprofit organization that brings the perspective of purchasers to APCD efforts. They also provide tools such as report cards on States' efforts on price transparency to help purchasers and other stakeholders understand issues related to payment reform and transparency.
- The **Health Care Cost Institute** is a nonprofit organization whose goal is to provide access to health care cost and utilization data to researchers and policymakers trying to understand the factors influencing health care costs. They create twice-yearly cost reports based on claims from four major insurers.
- The **Robert Wood Johnson Foundation** has also been instrumental in bringing together multiple stakeholders to improve the quality of health care. For example, their Aligning Forces for Quality initiative has resulted in public report cards about quality in some States.
- In terms of measurement science, **NQF** is a leader in endorsing and encouraging implementation of evidence-based, valid, reliable measures that are meaningful to stakeholders, including consumers.

Other major sources of measures for this report included the literature review, public reporting Web sites from APCDs, and other online reports of cost and quality. Individual organizations, such as the National Committee for Quality Assurance (NCQA), the Quality Alliance Steering Committee (QASC), and Bridges to Excellence, also had useful measures.

Major sources of measures. The table below gives a high-level overview of the key sources of measures identified through the environmental scan and literature review. The focus was on measures that leverage the strengths of APCDs, namely ambulatory measures, episode-of-care measures, and cost measures. This review formed the basis of the measure inventory for use with APCDs. Because numerous measures were identified, the first step of the measure inventory was to prioritize the measures.

Table 1. Number of measures or number of public reports for potential use with APCDs

Source	Number
Literature review	65 papers*
NQF Administrative Claims measures	
Ambulatory quality	143 measures
Resource	9 measures
National Quality Measures Clearinghouse	
Episode measures	141 measures
Cost measures for physicians	74 measures
APCD public reports	7 Public Reporting Websites
Other public reports with cost or resource measures†	7 Public Reporting Websites or Reports
Other measure stewards or resources	
NCQA Relative Resource Use Measures	5 measures
Quality Alliance Steering Committee	22 measures
APCD Showcase	41 reports
Bridges to Excellence	4 NQF-endorsed measures
HealthPartners	2 NQF-endorsed measures

NQF: National Quality Forum; APCD: All-payer claims database; NCQA: National Committee on Quality Assurance.

*These papers provide measures or potential measures by describing one measure that is specific to the study question, using claims data, or describing the use of a group of measures that are already in use and are described elsewhere (e.g., NQF-endorsed measures, CMS measures).

†From a list compiled in AHRQ’s Evidence-based Practice Center Technical Brief Protocol. Public Reporting of Cost Measures in Health. <https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1838>.

The literature provides studies using individual measures, some with well-described technical specifications. NQF and NQMC are measure aggregators and provide access to structured technical specifications. Public reporting Web sites either explicitly use State APCD data or use a combination of several data sources and measures. Some of these can be adapted for use with APCD data, although technical specifications are not always easily available through the online resource. Lastly, several organizations contribute discrete groups of claims-based measures. For example, NCQA, QASC, and Bridges to Excellence have made the technical specifications publicly available or available on request.

Key categories and domains for measures that leverage APCD data. The team identified key concepts and measure categories described in the literature review and environmental scan. These key categories and domains were used to organize the measure inventory, as well as in prioritizing certain categories of measures.

For instance, the purpose of performance reporting (e.g., for choice, negotiation, or accountability) can help guide measurement choice in the following way: if the purpose of a public report is to support consumer choice of providers, then a cost measure that only shows the average insurance

reimbursement rate without including the patient out-of-pocket cost for each provider will not be helpful. However, if the purpose of the measure is to assist in negotiations between insurers and providers, average insurance reimbursement would be potentially more useful.

In addition, measures may be used for population health and policy purposes. For example, population-level measures of utilization and cost are important to strategic planning to help eliminate health care disparities at the State, regional, and local levels. APCD data may also be used in State operations such as budgeting and rate review.

Barriers to using APCDs for measurement and potential solutions to overcome them.

Although APCDs are meant to contain comprehensive claims data across settings and time, there are still many barriers to using APCDs for measurement. Issues with data completeness, quality, standardization, and access hamper such efforts. Identifying and resolving these barriers is critical to using APCDs to improve health care value. The table below summarizes key barriers to APCD data collection and use and potential solutions for overcoming them.

Table 2. Key barriers and potential solutions related to data availability, quality, and access

Barrier	Potential Solution
Missing data elements	<ul style="list-style-type: none"> • Develop separate submission mechanisms for certain information (e.g., plan benefit design elements). • Add fields that indicate non-claims-based information (e.g., medical home or capitated arrangements). • Develop public report card for completeness of data submissions by payer.
Low data quality	<ul style="list-style-type: none"> • Develop public report card for data quality. • Improve data infrastructure.
Lack of data standardization	<ul style="list-style-type: none"> • Establish industry standards (e.g., ANSI ASC X12 and National Council for Prescription Drug Programs). • Establish standard reporting frameworks.
Difficulty with data linkage and aggregation	<ul style="list-style-type: none"> • Establish Master Patient Index and Master Provider Index. • Consistently use National Provider Identification numbers.
Lack of data access and availability	<ul style="list-style-type: none"> • Improve data reporting.
Policy barriers and resource limitations	<ul style="list-style-type: none"> • Diversify funding sources. • Work toward harmonization of stakeholder interests.

Methodological issues or barriers pertaining to using APCD data for measurement. As more States develop APCDs and public reporting Web sites based on APCD data, it is critical to find valid measures that are relevant to stakeholders and feasible to implement on a large scale. Multiple methodological issues have arisen in pursuit of this goal. Key issues include:

- Inadequate measurement science that may threaten the validity and reliability of measures.
- Lack of standardization of measure concepts and specifications.
- Difficulties in implementing measures due to privacy concerns, denominator deficiency, difficult or inaccurate provider attribution, inadequate risk adjustment, or provider reluctance to participate in public reporting initiatives.
- Measure gaps, including methodological gaps and gaps in existing measures.

These methodological issues and measure gaps will need to be addressed for States and others to fully realize the potential of APCDs in increasing health care value. Some States are using their APCDs initially for public reports of State-level performance and substate (e.g., county or ZIP Code level) variations before the release of more granular analysis (e.g., by provider or payer). This allows early public reporting of policy-relevant data from APCDs for State decisionmakers, while the barriers described above are being addressed.

Conclusion. While it is clear that there is much to be learned to maximize the potential of APCDs and to reduce the difficulty and cost of using them, there is also national momentum building behind developing measures to be used with APCDs and defining the business cases for maintaining APCDs. This environmental scan provides an overview of both the potential for APCDs to generate the information needed to improve care, as well as caveats to keep in mind while doing so.

Measure Inventory

The overall goal of the measure inventory was to provide a useful and usable inventory of measures that can be derived from APCDs. The inventory is based on a literature review and environmental scan completed in January 2015. It consists of measures identified in spring 2015 from the environmental scan, in consultation with a multi-stakeholder TEP, and with AHRQ program officer input.

Collected measures related to cost, utilization, and quality focused on measures relevant to ambulatory care or measures across settings (inpatient and outpatient), because APCDs are distinguished from other widely available datasets (e.g., hospital discharge data) by the inclusion of claims across multiple settings. The measures also focused on high-priority conditions, chosen based on a brief environmental scan (spring 2015) of priority conditions identified by the Institute of Medicine, the National Quality Strategy, the Centers for Medicare & Medicaid Services, Healthy People 2020, and others (noted in Appendix B).

The high-priority areas included:

- Measures that are not condition specific, such as imaging and medication management, and
- Measures that address specific high-priority conditions, (i.e., cardiac disease, preventive services, kidney or bladder conditions, mental health and substance abuse diagnoses, diabetes, and gastrointestinal disorders).

Several use cases for APCD-based measures have been proposed, including:

1. Choice, defined as measures that inform consumer or purchaser decisions;
2. Negotiation, such as use of measures in price or contract negotiation;
3. Accountability, or measuring provider or plan quality and efficiency of care; and
4. Population health and policy, to assess population health and inform and assist in policy decisions for States (e.g., efforts to assess health reform efforts or decrease cost).

In the literature review, environmental scan, and discussions with the TEP, the team assessed the use case that might best support the development and use of APCDs and be most useful to the States sponsoring them. The TEP found that the State-based use case focused on population health management was likely the most immediately useful and feasible one.

The inventory includes 302 measures. Provided with each measure is a brief description of the measure, measure steward, NQF endorsement status, type of measure (cost, quality, or utilization), risk adjustment, and Web site or reference (if measure is from the literature) where the measure was found.

In addition, for a smaller subset of measures, the team conducted a “deep dive.” The team chose deep dive measures from high-priority topics with measures for cost, quality, and utilization. These measures contain more detailed measure specifications, information on whether validity and reliability testing were conducted, and information on use in Federal programs, among other details.

In summary, a large number of measures are available for potential use with APCDs, covering a broad range of topics. Measures were grouped together to assess cost, quality, and utilization for specific topic areas for an assessment of population health management within or across diseases.

Although there are some limitations to both APCD data and the measures themselves, the measure inventory demonstrates that APCDs are potentially powerful new tools for monitoring population health. They can be used to paint a more complete picture of health care delivery, across payers and across settings, in ways that have not previously been possible. With continued development of both APCDs and measurement, stakeholders such as States, payers, providers, and consumers can look to use the APCDs to help fulfill the Triple Aim of better health, better quality, and lower costs.

Data Evaluation of All-Payer Claims Databases

To further assess the potential and current limitations of APCDs to support health care measures, the team acquired APCD data from three States, along with accompanying documentation and online documentation. These data were used to conduct preliminary analyses.¹ The focus of the assessment was on tables and fields particularly relevant to the calculation of health care measures, such as those identified in the measure inventory.

Specifically, each APCD was reviewed to:

- Determine whether the fields received were correct (e.g., fields provided matched the documentation, values fell within expected ranges);
- Document any anomalies;

¹ In agreement with the States providing the data, the States are not named in this report. Resource limitations allowed limited analyses, focused primarily on descriptive statistics and missingness of data across the full, unmanipulated datasets. The results are presented in qualitative form in accordance with the project-specific Data Use Agreement for each APCD.

- Note any differences over quarters, across States, or both;
- Provide a preliminary assessment of whether the data could be used to construct reliable and valid health care measures; and
- Provide insight into future analyses that will further validate these data for use in measure construction.

Methods

The team analyzed data for each State separately and did not link the data across States. The APCDs evaluated were limited or restricted datasets to provide a fuller range of variables that could be used for measure construction and to provide information on potentially useful variables for validating APCD data elements in future projects. When feasible, the datasets include the majority of non-patient-identifying fields (except for unique, encrypted patient identifier) from medical claims, enrollment records, and provider records.

This initial evaluation focused only on medical and pharmacy claims. While some States include dental claims, these were not evaluated. Also, due to limitations in data availability and time available to obtain data, APCDs varied with regard to the inclusion of Medicare and Medicaid records, which limited the ability to compare analyses across States.

The analyses were conducted and are reported in accordance with each State's Data Use Agreement. Detailed data analyses were available to AHRQ and researchers to construct recommendations for future work.

The analyses were intended as a high-level overview of the existing, unmanipulated APCDs. The overall data assessment sought to assess:

1. Accessibility of data.
2. Basic usability of data and documentation received.
3. Availability of data elements needed to construct health care measures in the utilization, quality, and cost domains, such as unique identifiers, provider or facility identifiers, service dates, sociodemographic fields, payment fields, revenue codes, clinical diagnosis and procedure codes, payer, and claim status.
4. Consistency of the distribution of key data elements with the documented data standards and face validity of the values.

For all key fields, the percentage missing/populated, distribution of values for continuous fields, and frequencies for categorical variables (up to 50 most frequent values) were calculated. The team also assessed the number of unique patients, unique claims, and number of transactions by month. These numbers were compared with State populations obtained from the U.S. Census Web site for qualitative assessment only.

Results of Data Evaluation

Availability of APCD data. This report provides a brief summary of the availability of datasets. Because the project required a large number of potential variables to conduct the global assessment, the experience of obtaining data for this project may not reflect the experience of

obtaining data for other measurement efforts. Although up to 14 States had APCDs in 2014, only 7 had made those data available for extramural analyses. Four State APCDs were pursued with the intention of obtaining three for analysis.

The applications for data varied somewhat but were all reasonable in length and detail and were consistent with other data request applications. One application for a Limited Data Set version was denied due to the lack of consistency of this project's goals with the data release guidelines for that APCD. In some cases, the process for obtaining data required iterative submissions and communication with the APCD organizations.

Overall, the application process required more time than anticipated, but all remaining applications were approved, except for Medicaid data for one State. The data were provided at a cost consistent with other public datasets, and educational or government rates were frequently available. Some States provide data without charge to government or educational institutions.

Data usability assessment. The first step was to confirm that the expected data files were received with documentation and to assess the ability to convert the flat data files into SAS files. The analysts completing these tasks were experienced in taking in raw data of this type, such that the usability assessment here may have limited application to users with less data experience.

All data dictionaries, control totals, and user guides were received as they were available with the data. Few issues were experienced in loading the data. Upon data intake, the team learned some files were missing from one State dataset as specified in the documentation and some documentation did not match the data layout provided. One file could not be imported into SAS for unclear reasons. The APCD team worked with the project team to provide a file that could be read. When available, the record numbers reported in the control tables were observed in all resulting SAS files.

In general, States provided reasonable assistance with all issues identified. However, because APCDs are often run with limited staff, adequate time allocations were required to be made to overcome the limited staffing of these organizations.

Examination of the reference tables (e.g., lookup tables providing labels for market, diagnosis codes, etc.) for all three States did not show any problems with the values in these tables with regard to duplicates or values/contents in the files. One can use the lookup tables with confidence.

Evaluation of data elements and completeness. The team checked the databases for unique members, data elements used to construct health care measures, and other criteria such as data elements used to identify unique encounters.

Unique members in database. The team assessed the total number of unique members by month to observe any large fluctuations that may suggest concerns about the unique patient

identifier or completeness of the claims provided. In some cases, modest variability over time was found, which appeared from Web site documentation to be due to incomplete claim submission.ⁱⁱ

It was observed in one case that the total number of unique individuals could exceed the population within a State, suggesting incomplete matching of claims to individuals. This situation can occur due to carve-outs, multiple coverage for one individual, duplicate claims, or failures in the patient identifier encryption methods. Despite these two issues, which were addressed in subsequent data releases or clarifications, these analyses showed reasonable numbers of unique patients relative to State populations and consistency over time.

Data elements used to construct health care measures. APCDs are typically organized into eligibility, claim, and provider files. However, data elements across these files are important to construct health care measures of utilization, quality, and cost. This section discusses several groups of measures that serve unique functions in health care measures. For any given measure, a subset of these elements would be used.

This section aims to describe the completeness of the data in qualitative terms, consistent with the project-specific Data Use Agreements and to provide context for the findings of the limited analytic review. In many cases, further analyses would better evaluate the validity and reliability of these data elements. This section also notes similarities and differences between the definition and coding of the data elements, as such consistency across States provides additional measurement opportunities.

Data elements used to identify unique encounters and link encounters. Most measures derived using APCD data will at minimum require the distinction between the type of health care encounters represented by the claim, such as isolating outpatient visits or prescriptions. Here, the team made the assessment on the following:

- Unique patient identifiers,
- Service dates,
- Variables that provide clues to movement through the health care system (e.g., discharge status or admission source),
- Variables used to identify the type of encounter (e.g., hospital, pharmacy, outpatient), and
- Claim status, used to determine whether a claim is the final paid claim.

The analyses show the percentage missing on all claims for variables used to isolate or link encounters. Some variables are not expected to be 100 percent populated, because they are only used for specific claim types, such as inpatient claims. Other variables, while being fully populated, have important caveats.

Identifying unique encounters requires that data elements facilitate removal of any duplicate claims. Unlike administrative data, such as hospital discharge data, a single encounter often results in multiple claims from providers, facilities, and pharmacy and may also have multiple claims representing different claim processing statuses. Variables such as service dates and

ⁱⁱ Missing claims are frequently added in subsequent releases of the data.

bill/service types are essential to build a dataset with only unique encounters, but additional variables such as clinical variables or provider variables are also important. These variables are described later, because they also serve other functions in health care measurement.

Some measures count encounters of a certain type without concern for multiple unique encounters for a single patient, such as the total number of prescriptions of a class of medication per population. In this case, whether 50 or 500 people obtained the observed prescriptions is not relevant to the measure. These measures would require information about (1) the service or bill type, to isolate the encounter type of interest; (2) the service provider and date, to ensure unique claims; and (3) the status of the claim, to ascertain whether the claim is the final adjudicated claim. In this case, encrypted member or patient identifiers will also assist in identifying unique encounters, even if the measure does not require linking across encounters.

Measures designed to track care across the health care system, to track encounters for the same individual over time, or to construct “episodes of care” all must have robust patient identifying information. For this purpose, a unique encrypted patient identifier is necessary. Additional information about the patient, such as age, gender, date of birth, or clinical information about the encounter (e.g., diagnosis or procedure codes), can be used to ensure proper linkages across claims and to clean data.

Although the analyses did not specifically examine the quality of the patient identifying variable, some modest fluctuation over time was observed (i.e., more unique individuals were identified within a database than anticipated). This suggests that single individuals may have more than one unique identifier, complicating linkage of claims.

Because some variables are unique to inpatient encounters, the analyses were inconclusive but suggest further investigation to understand these variables. The completeness of inpatient-specific variables varied even within APCDs by 5 to 15 percent on the low end; in some cases, one or two variables appeared to be populated at a much higher rate. This variation suggests that the variables may contain either invalid or placeholder entries or may be used for claims of different service or facility types. Overall, the data completeness was robust.

Although users should always analyze for and potentially remove duplicate claims for the same encounter, many APCDs intake data from payers and prepare the data to assess quality and remove certain duplicate claims. For instance, in Maine, the State data steward applies an algorithm to remove duplicate pharmacy claims so that the same prescription will not be double counted.⁶

Some data cleaning undertaken by the APCD or the user will remove real multiple encounters that occur on the same day, such as multiple emergency department visits or the same drug administered twice on the same day. As of yet, however, few variables are available to assist with such targeted data cleaning. Further, the issue of duplicate claims likely far outweighs the claims lost.

Table 3. Select variables useful for identifying and linking encounters

Variable Type	Range of Nonmissing Across APCDs and Quarters	Notes on Expectations and Validity*
Unique patient identifier	100% consistent across quarters and years	Unique patient identifier is typically created and fully populated. Usefulness of the identifier may be affected by methods used to create identifier. Observed higher numbers than anticipated given known covered lives, suggesting multiple identifiers per patient.
Service date	100% consistent across quarters and years	Typically includes day, month, and year.
Admission date	Highly variable, 10%-100%	Required only for inpatient claims, so lower rates would be expected. Unclear whether quarters with 100% of claims with nonmissing admission dates represent valid dates or whether dates match the service date. In general, inpatient claim variables are not reported at the same rate in data, suggesting use of fields for non-inpatient claims or missing values.
Admission type	Highly variable, 16%-75%	Required only for inpatient claims, so lower rates would be expected. Unclear whether quarters with 75% of claims with nonmissing admission type represent valid data. In general, inpatient claim variables are not reported at the same rate in data, suggesting use of fields for non-inpatient claims or missing values.
Discharge date	Highly variable, 21%-93%	Required only for inpatient claims, so lower rates would be expected. Unclear whether quarters with 93% of claims with non-missing dates represent valid dates or whether dates match the service date. In general, inpatient admission date and discharge date are reported at different rates, suggesting use of fields for non-inpatient claims or missing values.
Discharge status	14%-37%	Required only for inpatient claims, so lower rates would be expected. In general, APCDs were less variable in reporting but reported frequently on more claims than anticipated (typically, 30% of claims are not inpatient).
Admission source	13%-41%	Required only for inpatient claims, so lower rates would be expected. In general, APCDs were less variable in reporting but reported frequently on more claims than anticipated (typically, 30% of claims are not inpatient).
Bill type, service type, billing provider type	95%-100%	APCDs have different ways of populating this key variable. National billing standards break out facility type from provider services. While all APCDs evaluated had the standard fields populated at similar rates, each had unique fields that were fully populated with the type of service or facility.
Claim status	100% consistent across quarters and years	Variable is reported consistently, but the available categories vary by APCD.

*Additional information on validity can be found in Discussion section of this report.

Data elements that describe the clinical aspects of an encounter. Because claims data focus on payment, the clinical information provided is limited. However, to provide context for and to justify a claim, certain useful clinical elements are included:

- International Classification of Diseases, Ninth Revision (ICD-9, and now, the Tenth Revision, ICD-10) codes for diagnoses and associated present on admission (POA) flags
- Current Procedure Terminology (CPT) or ICD-9/10 codes for procedures
- Drug codes for pharmacy claims and details about the prescription, such as whether it is a new prescription or refill, the quantity dispensed, and the anticipated day supply.

These fields were found to be nearly always populated as appropriate.

First listed or principal diagnosis codes were usually listed, while secondary diagnosis codes were listed in just over half of claims and further diagnoses dropped further. Although secondary diagnoses can be used across different types of claims, it is known that outpatient claims contain fewer diagnosis codes. Secondary diagnosis codes often do not affect reimbursement. But for inpatient claims, where these codes do affect reimbursement, the number of available diagnosis code fields can affect measures.

Even in the inpatient setting, the APCDs evaluated did not adequately support POA coding, which should be available for all secondary diagnoses (and can be assigned to principal diagnoses). The lack of secondary diagnosis codes and POA coding affects the usefulness of these data for adequate risk adjustment without linking across multiple encounter and pharmacy records to establish comorbidities for a patient.

Procedure or service codes could be assigned to office visits as well as inpatient visits, but high variability in the inclusion of these codes was observed. This finding is consistent with other similar data.

DRG data are of limited usefulness in all-payer data unless derived after data are received, as providers are often not required to report DRGs to payers that do not use DRG-based reimbursement.

Table 4. Variables that describe the clinical aspects of the encounter

Variable Type	Claims Populated Range in Quarter and State*	Notes on Expectations
Principal or first listed diagnosis	99%-100%	ICD-9 or ICD-10-CM codes
Secondary diagnoses	The secondary diagnosis codes range from 53%-66% for the second position, then appropriately drop in frequency after about 5 diagnosis codes.	Secondary diagnosis codes can be included in records of different types but typically are more frequent in inpatient records where secondary codes can affect reimbursement. Outpatient records often include a limited number of diagnosis codes.
POA	0%-33%	Typically accompany inpatient diagnoses only but variable not populated or reported for majority of States/quarters.

Variable Type	Claims Populated Range in Quarter and State*	Notes on Expectations
Admitting diagnosis	4%-31%	Most State-quarter combinations have high missing rates (more than 90% missing). Admitting diagnosis differs from principal diagnosis, which is assigned at discharge. Low rates of admitting diagnosis are consistent with other data.
ICD-9 Principal Procedure	3%-22%	ICD-9/10 procedure codes are used most frequently in inpatient or high-intensity outpatient settings (some emergency departments, ambulatory surgery). Expected percentage is low, because not all receive procedures. Some APCDs combine procedure classifications into one variable (CPT, ICD-9).
CPT #1	20%-94%	Some APCDs combine procedure classifications into one variable (CPT, ICD-9). Outpatient encounters can almost always receive a CPT service code for an office visit.
Secondary procedure codes	<10%	APCDs vary on number of procedure fields available.
DRG	3%-17%	Typically only for inpatient claims. Not all payers use a DRG system. This will be populated at a low rate and for younger patients, lower than the inpatient rate. The DRG version is reported for a quarter of cases but may be auto-populated.
National Drug Code	100%	Included in the pharmacy claims, but reliability was not assessed. ⁷
Prescription details such as refills, day supply, quantity	100%	Included in the pharmacy claims, but reliability was not assessed.

* Unless specific to a pharmacy claim as noted, all missing values are based on the medical claim records only.

Data elements used in attribution to a provider or payer. APCD data can be used to calculate measures at different levels, including geographic (e.g., State), payer, facility, or provider. Because of the nested nature of the data, it is possible to build measures, but attribution remains a complex problem in most data sources.

Providers may use multiple Medicaid numbers or National Provider Identification numbers or provide care across different locations based on administrative and logistic features of their practice. Therefore, it can be difficult to attribute all encounters to the provider. The team observed a large number of providers in some quarter-State combinations, suggesting single providers using multiple provider IDs. The team did not evaluate whether the location of service was populated for most provider claims or if the population was consistent with the data submission guidelines. In addition, prescribers were not consistently identified in pharmacy records.

It is important to evaluate the provider data across the data tables. For instance, the master provider table may have limited information on specialty, while the claims table has this field populated as expected. Provider address and name variables can be used to further validate providers but can also vary by table.

The data also do not provide definitive markers for either assigned or utilized primary care providers, medical homes, or other constructs. This information may be estimated based on plurality of visits, payment arrangements suggesting medical homes, or other methods. A current NQF committee has developed a report on the attribution of care to providers and facilities.

Table 5. Variables used for measure attribution

Variable Type	Missing Range in Quarter and State	Notes on Expectations
Submitter	100%	Some records have submitters that differ from the payer.
Payer	100%	In some record types, payer may be reported less frequently.
Pharmacy location	74%-99%	ZIP Code
Billing provider (any unique identifier)	100%	The best variable for identifying billing providers is not clear. APCDs may populate Medicaid or Medicare numbers, National Provider Identifier, or unique provider codes. Some have multiple fields populated. Observed high numbers of providers in some cases, suggesting multiple identifiers for one provider.
Service provider city/ZIP Code	0%-100%	State is always provided, but more granular data varies depending on data request. Must justify more granular data when available.
Location of service	58%-100%	Location of service varies for State-quarter combinations. Applies to provider records.
Primary insurance indicator	96%-100%	Flag or identification of primary insurance.

Variables used to calculate cost. The variables associated with payment for services, which allow tracking and measuring of expenditure trends, were nearly always populated. These included payer/submitter, diagnosis, procedure, and all payment fields (copay, deductible, coinsurance, prepaid, and insurance payment). This finding is not surprising given that these data files are built from claims submitted to payers for reimbursement.

In relation to out-of-pocket health care costs, although all records have information on co-insurance and deductible, few records contain any information on the premium patients paid. Another source of missed costs is the exclusion of denied claims, which may, if not paid by a secondary insurer, result in out-of-pocket payment.

Other cost information related to secondary or other payers varied, such as co-insurance days or Medicare payments. However, because the total number of records where these fields would be relevant is not known, an assessment cannot be made regarding the observed variation in the completeness of these variables. Not all APCDs contain information on secondary payers.

Table 6. Variables used to calculate cost

Variable Type	Nonmissing Range in Quarter and State	Notes on Expectations
Copay amount	100%	In claims file
Deductible amount	100%	In claims file
Amount prepaid	100%	In claims file
Amount paid	100%	In claims file

Variable Type	Nonmissing Range in Quarter and State	Notes on Expectations
Coinsurance	100%	In claims file
Allowed amount	100%	Key variable for cost measures
Insurance type	100%	In eligibility file
Year and month of enrollment	100%	In eligibility file
Coverage flags (medical, pharmacy)	100%	In eligibility file

Demographic variables. While demographic variables beyond age and gender are rarely used to construct measures, they do have roles in assessing disparities and in some cases, where the application indicates, risk adjustment. The demographic variables in APCDs are limited. Age and gender are well populated while race and ethnicity are more variable. The validity of the race variable was not assessed, but this should be assessed further since it is unlikely that race is accurately reported on all records and not all individuals report a second race.

Other variables, such as insurance type or the member’s residence, can be used as a proxy for other demographic variables, although imperfect. The member address and ZIP Code on file may not always match the actual residence of the member.

Table 7. Demographic variables

Variable Type	Missing Range in Quarter and State	Notes on Expectations
Race, primary	19%-100%	Potentially over- and underreported. Most datasets do not contain 100% quality race data.
Race, secondary	17%-100%	Potentially over- and underreported. Most datasets do not contain 100% quality race data.
Ethnicity or Hispanic flag	74%-100%	Potentially over- and underreported. Most datasets do not contain 100% quality race data.
Insurance type	100%	Used as a proxy, but primary insurance type may not capture dual-covered or dual-eligible individuals.
Member city/county	98%-100%	Counties and cities can be heterogeneous, but data are available for socioeconomic status of county communities.
Member ZIP Code	99%-100%	
Sex	100%(nonmissing)	
Age	100%(nonmissing)	

Discussion

With nearly 18 States that have or are readying legislation for APCDs and with more than 30 States showing increasing interest in setting up APCDs, they are potentially a rich data source that captures unique data points and presents unique measurement opportunities. Unlike other data sources derived from a sliver of the health care system, APCDs allow more comprehensive evaluations of care across the health care system and the integration of cost of processes of care.²

These data include a wider range of individuals than other data sources; namely, they include a large proportion of covered encounters for commercially insured individuals and in some cases

Medicare and Medicaid populations.² Theoretically, information on individuals is not lost when switching providers or insurance coverage.

Because of these unique attributes, APCDs offer several unique opportunities for measurement:

1. Ambulatory care. Although Medicare fee-for-service utilization is widely studied across the inpatient and outpatient setting, data on all-payer ambulatory care are limited. APCDs can capture process measures occurring in the outpatient setting, such as appropriate ordering of laboratory testing or imaging, outpatient visits for chronic disease, and ambulatory care for adverse events following ambulatory surgery or hospitalization.^{8,9}
2. Pediatric care. This area has been particularly difficult to measure, because most children receive all of their care in outpatient settings. Pediatric measures such as immunizations, filling of prescriptions for asthma, and appropriate antibiotic prescriptions already are defined in other data sources.
3. Cross-cutting, episodic care. Measures that capture the spectrum of care can help us learn more about our health care system and effective care transitions. In addition, measures that capture care over an episode of several encounters are less susceptible to variation in the locations where patients seek or receive care. APCDs can capture followup care and care coordination and identify more entities involved in a patient's care.⁹
4. Population health. Because APCD data capture large portions of the population and are relatively inexpensive to collect, they are prime candidates for use in population health measures. Utilization measures are already used as proxies for population health, and APCDs can aid in expanding that view to different types of care, capturing routine care utilization and acute care that does not result in hospitalization. Some processes can also be captured in APCD data that are relevant to population health or the measures can be used to assess policy changes or disparities.^{5,8-10} From a resource use perspective, APCD-based measures can provide a window into the health care system, understanding the value of care and how health care costs are tied to the local practice patterns and population.⁹
5. Value measure suites. Although claims data have been tagged a source of health care price transparency, when cost data are paired with quality data, value measure suites can provide more nuanced information to both consumers and payers. These measure sets can be used to learn about more efficient ways to achieve better outcomes.

Limitations of APCD Data for Measurement

Missing populations. Although APCDs do include a large number of health care claims, some populations are omitted.^{2,8,9,11} If the data from these populations differ systematically from the data included in the APCD, the omission will result in biased measures. APCDs typically do not include:

- Uninsured patients;
- Some substance use, mental health, and HIV populations, depending on State regulations;
- Worker's compensation patients;
- Tricare or Veterans Affairs data;
- Federal Employee Health Benefit Plan patients; or
- Indian Health Service patients.⁹

A recent Supreme Court case (*Gobeille v. Liberty Mutual*) emphasized that certain plans are exempt from data submission requirements, although some will choose to submit anyway.¹¹ In this case, the Court found that the Federal Employee Retirement Income Security Act statute that preempts State legislation pertains primarily to self-insured health plans.¹² Because many employer-sponsored plans across the Nation are self-insured, this potential gap is being watched closely.

APCDs also have different rules regarding which patients should be included in the dataset.^{1,8,12} Some States require all care for patients residing in the State, while others require those employed within a State or those employed by a company within the State.¹ Because these are all slightly different definitions, users must be clear about the desired population for their use and clean the data appropriately.^{1,9}

Data quality. APCDs are prone to several data quality concerns. Claims data do not have details included in electronic health record data, including laboratory and biometric values. Most importantly, they typically do not contain patient outcomes.⁷ In this case, linking when possible to hospital administrative data records, electronic health records, or State mortality data add data.¹³⁻¹⁵ The data also may not capture bundled payments.⁸

Many of the critical data elements were well populated and were defined using uniform billing standards or other standardized data collection methods. However, some data elements for the validity or reliability of the populated values were not assessed. Two critical data elements in particular create data quality challenges.

First, the patient identifier is often based on a combination of static patient numbers (e.g., Social Security number, date of birth) and contract plan number.⁷ If patients switch from one plan to another, even if not switching providers, the unique identifier may change.⁷ In addition, if patients switch providers, employers, or claims from carve-out coverage, it may result in additional identifiers for the same individual.⁷ This can create difficulty linking encounters over months when plan changes are likely to take place (e.g., open enrollment, first of the fiscal or calendar year).

In a similar manner, there is evidence that payers and providers are not fully consolidated.⁷ In this case, administrative features may lead to a provider using different identifiers.

Delay in data versus data completeness. In the evaluation and on State APCD data Web sites, the struggle between data timeliness and completeness is highlighted. Some submitters fail to provide complete or cleaned data to the APCD in a timely manner, leading to claims being omitted. Therefore, one sees drops in utilization or data completeness until those data are received.

Although late data are eventually folded into the data, the delay in receipt of the full complement of claims can limit the usability of APCDs. In particular, cost is a factor that can change quickly. If APCDs are to provide cost transparency to consumers, the data must be available quickly. Because of this limitation, consumer use cases may be the most challenging use case for APCD measures.

Measure development gaps. Although hundreds of potential measures were identified that might be applied to APCDs, a few major measurement gaps exist. Specifically, measures of behavioral health, potentially due to data limitations, are not well developed or cannot be applied to the data.⁸ Pediatric measures focus primarily on a few clinical topics and could capture additional care.

Measure sets are likely to provide the best picture of health care value. Often, however, for the same population, few measures are available for cost, quality, and utilization.⁸ In particular, our review showed that quality measures were lacking. Again, data limitations may hamper the development of quality measures that assess outcomes using only APCD data, but linking to other datasets may allow stronger measures.

Finally, many measures were not rigorously tested for validity and reliability. Others may have been tested for reliability and validity for a specific level of analysis (e.g., health plan) but not another (e.g., physician group). Testing is required before a measure can be fully implemented.

Ongoing Efforts To Improve APCD Data

In speaking with States and reviewing State Web sites, the team identified a large number of ongoing activities related to improving data.^{10,12} These included:

- Ongoing evaluation of documentation and the release of validation reports by States,
- Work with researchers or other interested parties to assess the data against other data sources (e.g., mortality vital records) and
- Creation of reports based on APCD data.

Some States are isolating different methods for categorizing cost data that will best accommodate and present resource use information.⁸ Finally, some States have active user communities that provide an effective feedback loop on both usability and data quality issues to State organizations.

APCD core set of data elements. In 2009, the APCD Council and AHRQ began to develop a common core set of data elements that most of the APCD States were capturing. The result of that work was the APCD Core, a set of data elements common to most APCDs, which would provide a foundation for new States to grow their APCDs. Since then, the APCD Council has worked with two data standards maintenance organizations to develop standards based on electronic transactions used for claims adjudication. States have varied in their uptake of the APCD Core, with some adopting a unique format or, more commonly, using the core as a starting place for submission requirements.

In light of the 2015 *Gobeille v. Liberty Mutual* ruling, the APCD Council and the National Academy of State Health Policy have convened States to consider options for working with the U.S. Department of Labor (DOL), including development of a uniform dataset standard referred to as the common data layout (CDL). Although it remains in draft form, once finalized, the CDL can be referred to in work with DOL to identify a process for submitting self-funded plan data to State APCDs.

Federal interest in standards. Recognizing the emerging need for national standards, the National Committee on Vital and Health Statistics (NCVHS) convened a day-long public meeting, “Hearing on Claims-based Databases for Policy Development and Evaluation:- Overview and Emerging Issues,” in Washington, DC, on June 17, 2016. The purpose of this meeting was to highlight the current state of development, challenges, issues, and opportunities faced by claim-based databases, including State-based APCD initiatives and private-sector multipayer claim-based database efforts.

The meeting had the following priorities:

- Engage stakeholders on key issues related to claims-based databases.
- Identify priority areas and opportunities for recommendations to the Secretary of Health and Human Services and the health care industry.

Of particular interest to the Committee was to identify benefits, efficiencies, and barriers to the adoption of a common claim-based database and APCD reporting standard, to outline a roadmap for achieving standardization, and to determine how NCVHS could provide support.

NCVHS is currently reviewing the results of the meeting and formulating recommendations to the Secretary that are within the scope of the Committee’s statutory role and priorities identified at the June meeting.

Conclusion and Future Directions

APCDs hold promise as a way for policymakers, payers, providers, and consumers to gain information about health care quality, cost, and utilization and use it to help achieve higher quality, lower cost care. In addition, APCD measures can help us:

- Understand utilization patterns in national and local health care systems,
- Understand how patients move through the system, and
- Assess the variety of providers touching a patient during an episode of care to inform studies on the attribution of measures to providers.

Measurement using APCDs is still a developing field. However, across multiple phases we identified:

1. Increasing interest and activities aimed at improving and using APCD data;
2. Varied measures in high-priority clinical areas theoretically feasible using APCD data;
3. Well-populated key data elements, many of which were systematically coded between States; and
4. Ongoing national activities that aim to address some of the limitations of APCD data for measurement.

Future efforts can provide useful information and tools for using APCD in health care measurement. Specifically, mental health and pediatric care are areas where APCD data can uniquely contribute to population health. Measures that cut across the health care system and value-based measure suites should also be considered for future measure development.

This study was a preliminary assessment of APCDs as a source of measurement data. Additional and detailed data analyses are required before the full assessment of APCD-based measures can be completed. These analyses should first focus on key data elements stratified by claim type. This will require building a dataset by removing potential duplicates or claims that have been replaced by a final paid claim, if necessary.

Next, the consistency of data elements across measured entities, such as payers or providers, should be assessed to determine whether any consistent bias exists. This review is particularly important for diagnosis and procedure information, where the data are variable and differences between the claim data submitted may result in biased measures.

Finally, some select measures should be applied to the data to assess trends across measured entities and time, reliability of key data elements, and validity of measure scores. Aside from measure analyses, the continued efforts to improve documentation will provide a record of changes and improvements in the data. Benchmarking efforts will provide more context for data users and valuable public data.

With ongoing efforts, APCDs remain the most comprehensive source of data for monitoring population health and health care across the full system, assessing interventions aimed at improving population health, and assisting in assessing the value of health care.

References

1. APCD Council A; Porter J, Love D, Peters A, et al. The basics of all-payer claims databases: a primer for states. Princeton, NJ: Robert Wood Johnson Foundation; January 2014. http://www.rwjf.org/content/dam/farm/reports/issue_briefs/2014/rwjf409988. Accessed December 21, 2016.
2. Peters A, Sachs J, Porter J, et al. The value of all-payer claims databases to states. *N C Med J* 2014;75(3):211-3. <https://www.ncbi.nlm.nih.gov/pubmed/24830498>. Accessed December 21, 2016.
3. Pollack CE, Hussey PS, Rudin RS, et al. Measuring care continuity: a comparison of claims-based methods. *Med Care* 2016;54(5):e30-4. PMID:PMC4101051 [Available on 2017-05-01].
4. Price transparency: an essential building block for a high-value, sustainable health care system. Action Brief. Berkeley, CA: Catalyst for Payment Reform; November 2012. http://www.catalyzepaymentreform.org/images/documents/CPR_Action_Brief_Price_Transparency.pdf. Accessed December 21, 2016.
5. Miller P, Love D, Sullivan E, et al. All-payer claims databases: an overview for policymakers. Durham, NH: SCI, NAHDO, RAPHIC, APCD Council; May 2010. <http://www.apcdouncil.org/publication/all-payer-claims-databases-overview-policymakers>. Accessed December 21, 2016.
6. Topsham, ME: Maine Health Management Coalition; 2014.
7. Martin BC, Shewale A. Validity of electronic prescription claims records: a comparison of commercial insurance claims with pharmacy provider derived records. *Innov Pharm* 2014;5(1). <http://pubs.lib.umn.edu/innovations/vol5/iss4/7/>. Accessed December 21, 2016.
8. Block R. State models for health care cost measurement: a policy and operational framework. New York, NY: Milbank Memorial Fund; 2015. http://www.milbank.org/wp-content/uploads/2016/04/Milbank_Report-State_Models_for_Health_Care_Cost-2.pdf. Accessed December 21, 2016.
9. Ricards J, Blewett L. Making use of all-payer claims databases for health care reform evaluation. (Prepared by the State Health Access Data Assistance Center (SHADAC) at the University of Minnesota under Contract No. HHSP23320100024WI to the Office of the Assistant Secretary for Planning and Evaluation.) Minneapolis, MN: SHADAC; July 2014. <http://www.shadac.org/publications/making-use-all-payer-claims-databases-health-care-reform-evaluation>. Accessed December 21, 2016.
10. Miller P, Peters A, Porter J, et al. New York's all-payer database: a new lens for consumer transparency. Durham, NH: APCD Council and New York State Health Foundation; September 2015. <https://www.apcdouncil.org/publication/new-york%E2%80%99s-all-payer-database-new-lens-consumer-transparency>. Accessed December 21, 2016.
11. Key regulatory issues facing APCD states post *Gobeille v. Liberty Mutual*. Durham, NH: APCD Council and National Academy for State Health Policy; April 2016. <https://www.apcdouncil.org/publication/key-regulatory-issues-facing-apcd-states-post-gobeille-v-liberty-mutual>. Accessed December 21, 2016.
12. Miller P, Peters A. APCD legislation: review of current practices and critical elements. Durham, NH: APCD Council; November 2013. <https://www.apcdouncil.org/publication/apcd-legislation-review-current-practices-and-critical-elements>. Accessed December 21, 2016.
13. All payer claims databases: unlocking the potential. NEHI Issue Brief. Cambridge, MA: Network for Excellence in Health Innovation; December 2014. http://www.nehi.net/writable/publication_files/file/apcd_brief_nehi_draft_1210_final.pdf. Accessed December 21, 2016.
14. Love D, Steiner C. Key State Health Care Databases for Improving Health Care Delivery. Salt Lake City, UT: National Association of Health Data Organizations; February 2011. https://www.nahdo.org/sites/nahdo.org/files/Resources/HDD_APCD_Fact_Sheet_021411.pdf. Accessed December 21, 2016.
15. Afghahi A, Mathur M, Thompson CA, et al. Use of gene expression profiling and chemotherapy in early-stage breast cancer: a study of linked electronic medical records, cancer registry data, and genomic data across two health care systems. *J Oncol Pract* 2016;12(6):E697-709. doi: 10.1200/JOP.2015.009803. Epub 2016 May 24. <http://ascopubs.org/doi/full/10.1200/JOP.2015.009803>. Accessed December 21, 2016.

Appendix A: Technical Expert Panel and Learning Network Members

Technical Expert Panel

Marilyn Schlein Kramer, M.B.A.
Vice President/General Manager
Provider and Health Plan Markets, Health
Data & Management Solutions, Inc.
Former Deputy Executive Director
Commonwealth of Massachusetts
Center for Health Information and Analysis
Boston, Massachusetts

Jonathan Mathieu, Ph.D.
Vice President for Research & Compliance
and Chief Economist
Center for Improving Value in Health Care
Denver, Colorado

David Newman, J.D., Ph.D.
Executive Director
Health Care Cost Institute
Washington, DC

Kenneth Park, M.D.
Vice President
Real World Evidence Solutions
IMS Health
Former Vice President
Payer and Provider Solutions
Anthem, HealthCore
Plymouth Meeting, Pennsylvania

Doris Peter, Ph.D.
Director

Health Ratings Center
Consumer Reports
New York, New York

Norman Thurston, M.A.
Director
Office of Health Care Statistics
Utah Department of Health
Salt Lake City, Utah

Rachel Werner, M.D., Ph.D.
Associate Professor of Medicine
University of Pennsylvania
Senior Fellow, Leonard Davis Institute of
Health Economics
Core Faculty
Center for Health Equity Research and
Promotion
Veterans Affairs Medical Center
Philadelphia, Pennsylvania

Learning Network Members

APCD Council
National Association of Health Data
Organizations

Appendix B: Prioritization of Conditions for Measure Inventory

Organizations from which high-priority condition lists were reviewed

Source	Full Title	Link to Source
American Academy of Pediatrics	Priority Topics	No link available
Center for Medicare & Medicaid Innovation	Priority Measures for Monitoring and Evaluation	No link available
Health Partners	Health Partners National Quality Forum (NQF)-endorsed measures	http://www.qualityforum.org/QPS/QPSTool.aspx Search Term: Health Partners
Health Resources and Services Administration (HRSA)	HRSA Clinical Quality Measures	https://www.hrsa.gov/quality/toolbox/measures/index.html
Healthy People 2020	Healthy People 2020 Leading Health Indicator Topics	https://www.healthypeople.gov/2020/leading-health-indicators/2020-LHI-Topics
Institute of Medicine	100 Initial Priority Topics for Comparative Effectiveness Research	http://www.nationalacademies.org/hmd/~/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/Standard%20Alone%20List%20of%20100%20CER%20Priorities%20-%20for%20web.ashx
Medical Expenditure Panel Survey (MEPS)	MEPS Topics: Priority Conditions--General	http://meps.ahrq.gov/data_stats/MEPS_topics.jsp?topicid=41Z-1
National Committee on Quality Assurance (NCQA)	Bridges to Excellence NQF-endorsed measures	http://www.qualityforum.org/QPS/QPSTool.aspx Search Term: Bridges to Excellence
NCQA	NCQA Relative Resource Use Measures – NQF endorsed	http://www.qualityforum.org/QPS/QPSTool.aspx Search Term: Relative Resource Use Measures
NQF	NQF 2012 Multiple Chronic Conditions Measurement Framework, Appendixes B and C	http://www.qualityforum.org/Projects/Multiple_Chronic_Conditions_Measurement_Framework.aspx
National Quality Strategy	Annual Progress Report to Congress: National Strategy for Quality Improvement in Health Care 2013 Report, Appendixes A and B 2014 Report, Figure	2013 Report: http://www.ahrq.gov/workingforquality/nqs/nqs2013annlrpt.htm 2014 Report: http://www.ahrq.gov/workingforquality/reports/annual-reports/nqs2014annlrpt.htm
National Quality Strategy	Input to the Secretary of Health and Human Services on Priorities for the National Quality Strategy, 2011	http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=68238
Office of the Assistant Secretary for Health (OASH)	OASH List of Chronic Conditions	https://www.cdc.gov/pcd/issues/2013/12_0239.htm#table3_down
Patients Like Me	Patients Like Me: Conditions	https://www.patientslikeme.com/conditions
Quality Alliance Steering Committee	High-Value Health Care Project program results report, Appendix 7	http://www.rwjf.org/content/dam/farm/reports/program_results_reports/2011/rwjf71110