

October 2020

Federal Policy Options to Realize the Potential of APCDs

Matthew Fiedler and Christen Linke Young

USC-Brookings Schaeffer Initiative for Health Policy

This report is available online at: <https://www.brookings.edu/research/federal-policy-options-to-realize-the-potential-of-apcds/>

USC Schaeffer
Leonard D. Schaeffer Center
for Health Policy & Economics

B | Economic Studies
at BROOKINGS

EDITOR'S NOTE

This white paper is part of the USC-Brookings Schaeffer Initiative for Health Policy, which is a partnership between the Economic Studies Program at Brookings and the USC Schaeffer Center for Health Policy & Economics. The Initiative aims to inform the national health care debate with rigorous, evidence-based analysis leading to practical recommendations using the collaborative strengths of USC and Brookings. The Commonwealth Fund provided a grant to the Brookings Institution to support the writing of this report.

ACKNOWLEDGEMENTS

The authors thank Kathleen Hannick and Spoorthi Kamepalli for excellent research assistance and Brianna Nicker for excellent editorial assistance.

Executive Summary

An all-payer claims database (APCD) is a system that collects health care claims and related data from all (or nearly all) entities that pay for health care services in a geographic area, including private and public health plans. Today, 23 states have APCDs, and they are valuable tools that virtually anyone with a stake in the health care system—including consumers, employers, health care providers, health insurers, researchers, and policymakers—can use to better understand the system and find ways to improve it. Indeed, in the states that have them, APCDs can provide a comprehensive picture of health care spending, health care delivery, and health insurance enrollment, and they offer insights that no other data can replicate, particularly with respect to private health insurance markets.

Unfortunately, today's APCDs have important limitations. Most prominently, a 2016 Supreme Court decision, *Gobeille v. Liberty Mutual Insurance Co.*, held that states may not require data collection from non-governmental self-insured group health plans. Because self-insured plans represent 61% of enrollment in employer coverage—and about one-third of all covered people—this decision left a large gap in state APCDs.

Current APCDs also have limitations that predate the *Gobeille* decision and arise from the fact that they are state entities that operate independently of one another. Because each state has its own data submission protocols and data access procedures, combining data from multiple states is challenging, which can frustrate efforts to compare experiences in different states. Further, because APCDs are state-controlled, federal policymakers have no automatic access to APCD data. Relying on a patchwork of state APCDs also forfeits potential economies of scale, which increases administrative costs for both payers and APCDs themselves and may prevent APCDs from making valuable investments in public reporting or data quality. Further, half the U.S. population lives in a state without an APCD.

Federal policymakers have multiple options to address the limitations of current APCDs:

- **Enable state collection of self-insured data:** Federal policymakers can directly restore state APCDs' ability to collect data from self-insured plans. Congress could enact simple legislation that would achieve this objective. Alternatively, the Supreme Court's decision in *Gobeille* described a legal pathway by which the federal government could authorize states to collect these data on its behalf without new legislation. In either case, policymakers could choose to limit the ability to collect data from self-insured plans to state APCDs that meet certain conditions, like collecting data in a uniform format to reduce administrative burden on payers or providing adequate researcher access to the data.
- **Create a national APCD:** The federal government could also build a national APCD that would collect data from all payers in all states. The federal government and outside researchers could then access this rich national dataset, subject to appropriate privacy safeguards. Data from a national APCD could also be shared with state APCDs, which might need to adapt their systems to accept it, but would then be freed from the burden of collecting data themselves and could focus their limited resources on high-value efforts to support broader use of the data.

In designing such a system, policymakers would need to pay careful attention to privacy and data security concerns. A national APCD could adopt security procedures like those that govern existing federal health care databases. It could also implement data access procedures similar to those that govern Medicare claims data today; in particular, it could bar users from disclosing anything other than aggregated results, require them to abide by rigorous data security practices, limit their access to only those data elements required for their planned analyses, and require them to access and analyze data in a secure computing environment controlled by the federal government. Policymakers could also bar an APCD from holding direct individual identifiers, although such restrictions would need to be crafted carefully to avoid reducing the APCD's capabilities.

The federal government already has the authority to collect these data and could do so without new legislation, although such an initiative is more likely to be pursued and completed if Congress directs and funds the work. Policymakers could also consider housing this initiative within a non-profit, rather than a federal agency, as proposed in legislation recently considered in the Senate Health, Education, Labor, and Pensions Committee, although such an approach poses some governance concerns and would likely require new legislation.

- **Expand state APCD coverage and harmonize state APCDs:** Policymakers could also pursue a hybrid approach that would expand, improve, and harmonize the existing network of state APCDs. In this model, policymakers would take three steps: (1) authorize state APCDs to collect self-insured data; (2) provide grants to states to encourage creation of new APCDs; and (3) require state APCDs to collect data through a standardized process and share their data with a federal “clearinghouse,” similar to the federal clearinghouse that currently exists for state hospital encounter data. The clearinghouse could support research projects that require data from multiple states and facilitate use of these data by federal policymakers, while the grants might encourage more states to create APCDs. However, some states would likely continue to lack APCDs under this approach, and each state APCD would still need to invest in its own infrastructure for collecting and maintaining claims data. The federal government likely already has the authority to pursue this approach, but as with creating a national APCD, such an initiative is more likely to be undertaken with Congressional support.

We believe that creating a national APCD is the best of these approaches. With a small investment of resources relative to total US health care spending, policymakers could create a tool that offers a comprehensive picture of the health care system that no existing data source can come close to matching, thereby accelerating efforts by a wide range of public and private actors to better understand and ultimately improve American health care. While efforts to expand the number of state APCDs and harmonize existing APCDs could also improve on the status quo (and the pre-*Gobeille* status quo) and might encounter somewhat less stakeholder opposition, such a project would achieve less than creating a truly national APCD, and we view it as a decidedly second-best alternative.

If the political will cannot be summoned to pursue these larger projects, either Congress or federal agencies should act swiftly to at least restore states’ ability to collect self-insured data. While not cost free, this is a fairly simple undertaking that will make existing APCDs more useful and comprehensive. It would allow state APCDs to maximize their own potential, hopefully building a constituency for creating a better national infrastructure over the longer-term.

Stakeholders might raise objections to the policy approaches we recommend here, but policymakers should not be deterred from moving forward. Payers might raise concerns about the burden of reporting data to APCDs, but since payers already provide similar data to many entities, including state APCDs, under the status quo, the incremental burden associated with these proposals would likely be modest. There is also an ongoing debate about whether disclosure of payers’ negotiated prices, including through an APCD, could put upward pressure on prices. While the evidence on this question is mixed, policymakers could prevent disclosure of negotiated prices if they wished, albeit not without reducing the utility of APCD data in important respects. Privacy and security concerns are also sometimes raised in response to options that involve the federal government holding claims data. However, APCDs do not appear to present any novel privacy or security issues relative to other government undertakings, and these concerns can be addressed using privacy and security standards similar to those that have proven successful in protecting other sensitive data, like Medicare claims data. We also note that some stakeholders sell data products that would become less valuable if APCD data became more robust and widely available, which may lead them to oppose APCD expansion, but this concern likely should not factor into policymakers’ decisions.

Contents

A Primer on APCDs1
What is an APCD?1
Uses of APCD Data 3
Stakeholder Objections to APCDs..... 5
Administrative Burden 5
Upward Pressure on Negotiated Prices..... 6
Privacy and Security 7
Data Uses 7
Unstated Stakeholder Objections..... 7
Limitations of Existing APCDs 8
The Gobeille Decision and its Consequences 8
The Court’s Decision in Gobeille v. Liberty Mutual 8
Consequences of the Gobeille Decision..... 9
Limitations of Relying on a Patchwork of State APCDs..... 10
Challenges in Combining Data from Multiple States..... 10
Incomplete Geographic Coverage 11
Inaccessibility to Federal Policymakers 11
Inability to Exploit Economies of Scale.....12
Federal Policy Options12
Enable State Collection of Self-Insured Data15
Legislative De-Preemption15
Agency De-Preemption.....16
Create a National APCD 18
Architecture of a National APCD.....19
Uses of APCD Data 20
Role of Existing State APCDs21
Privacy and Security Safeguards 22
Funding Requirements..... 24
An Alternative Governance Structure: Housing a National APCD Within a Non-Profit..... 25
Harmonize State APCDs and Create a Federal Clearinghouse for APCD data 26
A Note on “Federated” Alternatives to APCDs 28
Conclusion 29

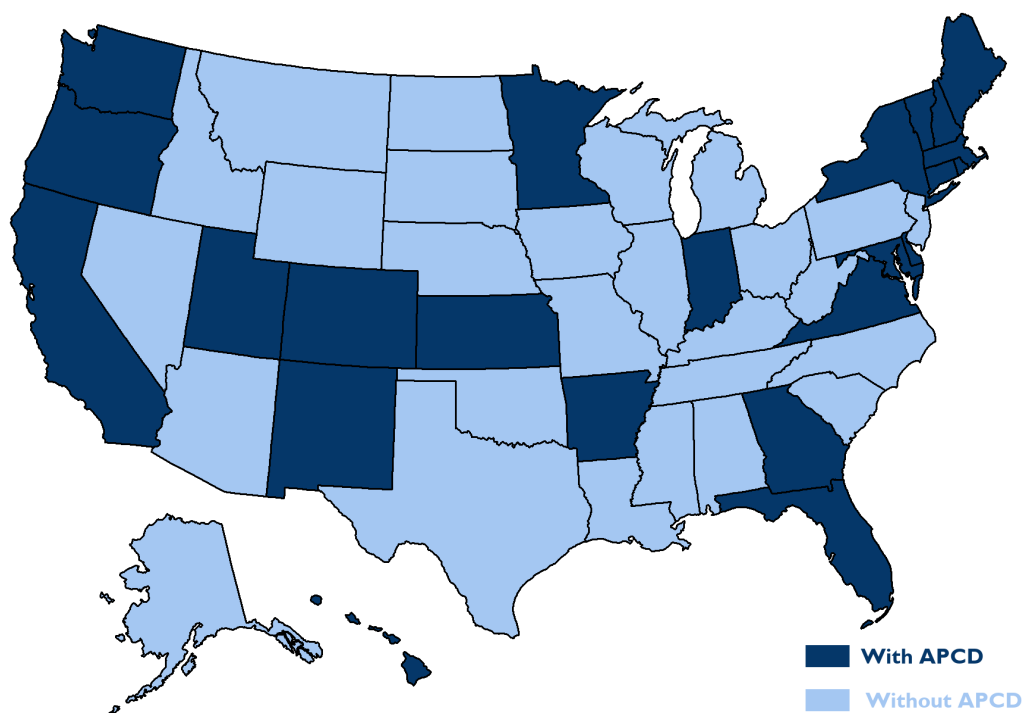
A Primer on APCDs

We begin this report by providing a brief overview of what an all-payer claims database (APCD) is, what APCDs can be used for, and the main objections raised against APCDs. Readers interested in more background on state APCDs may wish to refer to more comprehensive introductions elsewhere.¹

What is an APCD?

An APCD is a system that seeks to collect health care claims and related data from all (or nearly all) entities that pay for health care services in a geographic area, including private health insurance plans, Medicare, and Medicaid. All existing APCDs operate at the state level, and 23 states accounting for half the U.S. population currently have an APCD in operation or active implementation, as illustrated in Figure 1.² Several more states have APCD-like entities for which data submission is voluntary but which still reach a meaningful fraction of the state’s insurance market.³

Figure 1. States with APCDs, 2020



Note: California, Georgia, Hawaii, Indiana, and New Mexico APCDs are in implementation.
Source: Adapted from APCD Council

¹ Jo Porter, Denise Love, Ashley Peters, Jane Sachs, and Amy Costello, “The Basics of All-Payer Claims Databases: A Primer for States,” *All-Payer Claims Database Council*, January 2014, <https://www.apcdouncil.org/publication/basics-all-payer-claims-databases-primer-states>; Jo Porter, Denise Love, Amy Costello, Ashley Peters, Barbara Rudolph, “All-Payer Claims Database Development Manual: Establishing a Foundation for Health Care Transparency and Informed Decision Making,” *All-Payer Claims Database Council*, February 2015, <https://www.apcdouncil.org/manual>.

² For additional details on the data underlying Figure 1, see All-Payer Claims Database Council, “Interactive State Report Map,” <https://www.apcdouncil.org/state/map> (last visited October 19, 2020). The APCD Council reports that West Virginia began development of an APCD for which implementation has since stalled.

³ The states are Texas, Wisconsin, Michigan, Oklahoma, and South Carolina. All-Payer Claims Database Council, “Interactive State Report Map,” <https://www.apcdouncil.org/state/map> (last visited October 19, 2020).

Box 1: Glossary

- Group health plan. A health benefit plan offered by an employer to its employee.
- Insured group health plan. A product sold to an employer under which an insurance company assumes responsibility for paying enrollees' health care claims.
- Self-insured group health plan. A group health plan in which the employer is directly responsible for paying its employees' health care claims. Self-insured group health plans can (and generally do) hire an outside contractor, typically an insurance company, to handle the major administrative functions of the plan, including constructing provider networks and processing claims.
- Administrator. The entity who administers the benefits of a self-insured group health plan. Administrators are usually, but not always, a third-party under contract with the group health plan. Also called a third-party administrator when administration is conducted under a separate contract.

Most APCDs are directly operated by a state agency, but a few states delegate operation of their APCDs to a non-governmental entity.⁴ Even where an APCD is operated by a state agency, the work of collecting, cleaning, and maintaining data from payers is frequently contracted out to a vendor.⁵

In all states with APCDs, state law compels health insurers and the state's Medicaid program to submit data to the APCD. Most state APCDs also obtain Medicare data via agreements with the federal government. Importantly, as we discuss in much greater detail below, federal law has prevented states from placing similar requirements on most self-insured group health plans since 2016, and it appears that states are unable to collect data for most self-insured enrollees. Because self-insured plans represent about 61% of enrollment in employer coverage and about one-third of all people with coverage, the "all payer" label is something of a misnomer as applied to existing state APCDs.⁶ Further, state APCDs do not collect data from certain other payers, like the Federal Employee Health Benefits Program, TRICARE, or the Veterans Administration.

The core of an APCD is health care claims data. A health care claim is generated for each service an insurer pays for and contains a variety of useful information, including the type of service, the patient that received the service, the provider that delivered the service, the date of delivery, the diagnosis that precipitated the service, and—crucially—what the insurer paid for the service as well as what the enrollee paid in cost-sharing. Prescription drug claims contain similar information. APCDs generally also collect various related information held by payers that is useful for analytic purposes, which may include enrollee demographic characteristics like age and zip code, as well as characteristics of enrollees' coverage such as network characteristics and plan premiums.

⁴ For additional detail on each state APCD's governance structure, see All-Payer Claims Database Council, "Interactive State Report Map," <https://www.apcdouncil.org/state/map> (last visited October 19, 2020).

⁵ Jo Porter, Denise Love, Ashley Peters, Jane Sachs, and Amy Costello, "The Basics of All-Payer Claims Databases: A Primer for States," *All-Payer Claims Database Council*, January 2014, <https://www.apcdouncil.org/publication/basics-all-payer-claims-databases-primer-states>; Jo Porter, Denise Love, Amy Costello, Ashley Peters, Barbara Rudolph, "All-Payer Claims Database Development Manual: Establishing a Foundation for Health Care Transparency and Informed Decision Making," *All-Payer Claims Database Council*, February 2015, <https://www.apcdouncil.org/manual>.

⁶ For estimates of the share of employer market enrollment in self-insured plan see "2019 Employer Health Benefits Survey," *Kaiser Family Foundation*, September 25, 2019, <https://www.kff.org/report-section/ehbs-2019-section-10-plan-funding/>. The share of total enrollment in self-insured plans was calculated using the estimates of total employer market enrollment available at Kaiser Family Foundation, "Health Insurance Coverage of the Total Population," *Kaiser Family Foundation*, <https://www.kff.org/other/state-indicator/total-population/> (last visited September 15, 2020). In practice, states also generally exempt insurers with very limited enrollment from reporting to the APCDs. See Jo Porter, Denise Love, Ashley Peters, Jane Sachs, and Amy Costello, "The Basics of All-Payer Claims Databases: A Primer for States," *All-Payer Claims Database Council*, January 2014, <https://www.apcdouncil.org/publication/basics-all-payer-claims-databases-primer-states>.

APCDs fund themselves through a combination of mechanisms, including direct state funding, federal grant funding, and fees on data users.⁷ We are unaware of a data source that provides comprehensive information on APCDs' budgets, but three APCDs for which recent budget information is readily available have annual budgets ranging from \$1.6 million to \$4.4 million.⁸ Extrapolating these estimates nationwide implies that existing APCDs incur combined operating costs of less than \$100 million. For context, \$100 million amounts to less than 0.003% of national health expenditures in 2018 – or about \$1 per \$37,000 in health care spending.⁹ While APCD budgets do not include the costs that payers incur to report to APCDs, this figure suggests that if the uses of APCD data described in the next section facilitate even tiny reductions in health care spending or equivalent improvements in other aspects of health care system performance, then state investments in APCDs generate benefits that greatly exceed their costs.

Uses of APCD Data

Health care claims provide comprehensive information on what health care items and services (insured) people receive and how much is paid for those items and services. Consequently, claims can support many different types of analyses that have the potential to improve the health care system:¹⁰

- **Public reporting:** Many state APCDs use the data they hold to produce public reports on their states' health care systems. Some examine levels of and trends in health care utilization, spending, and quality in the state, which may be helpful to a variety of stakeholders, including individuals, the press, employers, providers, and insurers in understanding the current state of the health care system and making decisions related to it.¹¹ Others examine specific issues of current interest, such as opioid prescribing patterns, the effects of COVID-19 on the health care system, and the utilization of low-value services.¹² Some also use APCD data to create consumer-facing price transparency tools that allow patients to compare the prices charged by competing providers.¹³

⁷ Jo Porter, Denise Love, Ashley Peters, Jane Sachs, and Amy Costello, "The Basics of All-Payer Claims Databases: A Primer for States," *All-Payer Claims Database Council*, January 2014, <https://www.apcdouncil.org/publication/basics-all-payer-claims-databases-primer-states>.

⁸ "2019 Colorado All Payer Claims Database Annual Report", *Center for Improving Value in Health Care*, February 2020, <https://www.civhc.org/wp-content/uploads/2020/02/CO-APCD-Annual-Report-FY19.pdf>; "Washington All-Payer Claims Database 2019 Accomplishments," *Washington All-Payer Claims Database*, January 2020, https://www.ofm.wa.gov/sites/default/files/public/dataresearch/healthcare/pdf/2019_WA_APCD_accomplishments.pdf; "Collaborations in the Commonwealth, 2019 Annual Report & Strategic Plan Update," *Virginia Health Information*, http://www.vhi.org/About/annual_report.pdf (last visited September 17, 2020).

⁹ "National Health Expenditure Data," *Centers for Medicare & Medicaid Services*, December 17, 2019, <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData>.

¹⁰ For a review of a large number of applications of APCD data, see All Payer Claims Database, "APCD Showcase: States Leading by Example," <https://www.apcdshowcase.org/> (last visited September 17, 2020).

¹¹ "Annual Cost Trends Report," *Massachusetts Health Policy Commission*, <https://www.mass.gov/service-details/annual-cost-trends-report> (last visited September 17, 2020); Washington Office of Financial Management, "Washington State HEDIS Quality Measures (claims based) - Data Dashboard," <https://www.ofm.wa.gov/washington-data-research/health-care/health-care-access-utilization-and-quality/washington-state-hedis-quality-measures-claims-based-data-dashboard> (last visited September 17, 2020); "2018 Vermont Health Care Expenditure Analysis," *Green Mountain Care Board*, July 8, 2020, https://gmcbboard.vermont.gov/sites/gmcb/files/Misc/2018_VT_Health_Care_Expenditure_Analysis_Final_%20July_%2008_%202020.pdf.

¹² "2016 Virginia Low Value Services Report," *Virginia Health Information*, February 2017, <https://www.vhi.org/apcd/Virginia%20Low%20Value%20Services%20Report.pdf>; Utah Department of Health, "Preliminary COVID-19 Healthcare Trends: A Snapshot from Utah's All Payer Claims Database," August 25, 2020, <http://stats.health.utah.gov/latest-news/preliminary-covid-19-healthcare-trends/>; "Prescribing Opioids in Colorado," *Center for Improving Value in Health Care*, March 2019, <https://www.civhc.org/wp-content/uploads/2019/02/Opioid-Spot-Analysis-March-2019.pdf>.

¹³ Florida Health Price Finder, "Learn More About Getting The Most Out Of Your Health Care Dollars," <https://pricing.floridahealthfinder.gov/> (last visited September 17, 2020); New Hampshire Health Cost, "Compare Health Costs & Quality of Care," <https://nhhealthcost.nh.gov/> (last visited September 17, 2020).

- **Research:** APCDs can—and typically do—make the data they collect available to outside researchers, generally for a fee, with restrictions to protect patient privacy.¹⁴ State APCDs have supported research on a wide variety of topics, such as the effect of tiered networks on hospital choice and health care spending, the effect of Medicaid expansion on continuity of coverage, and the utilization of telehealth services.¹⁵
- **Direct policy applications:** APCD data can also directly support policy design and implementation. At the policy design stage, APCD data can give policymakers a better picture of their states' health care systems and help inform estimates of the potential consequences of policy changes. For example, New Hampshire used APCD data to better understand how the prices paid by commercial payers compared to the prices paid by its Medicaid program as part of an effort to redesign its Medicaid fee schedules, while Washington State used APCD data to inform the provider payment rate requirements under its new “public option.”¹⁶

APCD data also has applications in policy implementation. For example, APCD data has been used to determine out-of-network payment standards under state laws that address surprise billing.¹⁷ APCD data can also be used by state attorneys general to monitor for anti-competitive conduct in health care markets or evaluate proposed provider or insurer mergers.

APCDs have at least two important advantages over other claims data sources (or, at least, they would if they included data from all self-insured group health plans). First, APCDs can offer a more complete picture of the commercially insured population than other commercial claims databases, such as the databases maintained by IBM MarketScan, Optum (a subsidiary of United Healthcare), Blue Health Intelligence (a Blue Cross Blue Shield affiliate), and the non-profit Health Care Cost Institute (which, in its current incarnation, contains claims data contributed by Aetna, Humana, and United Healthcare, but will soon encompass data from Aetna, Humana, and Blue Health Intelligence). All of these databases capture a non-random subset of the commercial insurance market, and the subset of the market each database captures often changes over time; this introduces potential for bias in estimating marketwide averages and trends. This incomplete coverage, as well as limitations that contributing payers place on the use of their data, also limits the value of these databases for understanding how and why performance varies across different commercial payers and the factors that shape employers' and individuals' choices among competing insurance options.

Second, the fact that APCDs aspire to encompass all payers allows them to support analyses that would not be possible with databases that include only a subset of a state's insurance market. For example, APCDs can be used to study trends in insurance enrollment in the state's insurance market as a whole,

¹⁴ “Releasing APCD Data: How States Balance Privacy and Utility,” *All-Payer Claims Database*, March 2017, <https://www.apcdouncil.org/publication/releasing-apcd-data-how-states-balance-privacy-and-utility>.

¹⁵ Elena Prager, “Health Care Demand Under Simple Prices: Evidence From Tiered Hospital Networks,” *Northwestern University*, 2017, https://faculty.kellogg.northwestern.edu/models/faculty/m_download_document.php?id=523 <https://link.springer.com/article/10.1007/s11606-019-05101-8>;

Jiani Yu, Pamela J. Mink, Peter J. Huckfeldt, Stefan Gildemeister, and Jean M. Abraham, “Population-Level Estimates of Telemedicine Service Provision Using An All-Payer Claims Database,” *37 HEALTH AFFAIRS* 1931, December 2018, <https://www.healthaffairs.org/doi/abs/10.1377/hlthaff.2018.05116>.

¹⁶ Denise Love, William Custer, and Patrick Miller, “All-Payer Claims Databases: State Initiatives to Improve Health Care Transparency,” *The Commonwealth Fund*, September 2010, https://www.commonwealthfund.org/sites/default/files/documents/media_files_publications_issue_brief_2010_sep_1439_love_allpayer_claims_databases_ib_v2.pdf; Christina Cousart, “How Washington State Is Reducing Costs and Improving Coverage Value – A Q&A with its Health Benefit Exchange CEO,” *National Academy for State Health Policy*, August 5, 2019, <https://www.nashp.org/how-washington-state-is-reducing-costs-and-improving-coverage-value-a-qa-with-its-health-benefit-exchange-ceo/>.

¹⁷ Colorado Department of Regulatory Agencies, “Out-of-Network Health Care Provider Reimbursement,” <https://www.colorado.gov/pacific/dora/out-network-health-care-provider-reimbursement> (last visited September 17, 2020); Washington State Office of the Insurance Commissioner, “Arbitration and using the Balance Billing Protection Act Data Set,” <https://www.insurance.wa.gov/arbitration-and-using-balance-billing-protection-act-data-set> (last visited September 17, 2020).

potentially with far greater precision and granularity than survey data sources.¹⁸ They can also be used to study how often people transition among different types of coverage and what the consequences of those transitions may be for the cost and quality of patient care.¹⁹

APCDs are also a useful complement to state hospital encounter databases, which collect encounter records directly from hospitals.²⁰ Hospital encounter databases generally do not collect information on non-facility-based outpatient care or prescription drug utilization, nor do they collect information on the prices paid for health care services. Thus, they paint an incomplete picture of care patterns and are not suitable for analyzing health care spending. (On the other hand, encounter databases do capture utilization by uninsured people, which APCDs do not since APCDs collect data from insurers.)

Stakeholder Objections to APCDs

While APCDs have significant potential to inform the public, facilitate research, and directly support policymaking, they do have detractors. We discuss several common objections to APCDs, including that they create administrative burdens for payers, could place upward pressure on prices by disclosing confidential negotiated prices, could threaten privacy, or may be applied in an overly broad array of circumstances. We also briefly discuss concerns that often go unstated but may be important motivators of stakeholder opposition.

Administrative Burden

Payers frequently express concern that submitting data to APCDs is burdensome.²¹ Concerns about burden are raised particularly frequently by groups representing self-insured group health plans, who argue that being required to submit to state APCDs, as opposed to a single national APCD, burdens self-insured plans that have enrollees in multiple states.²² They often also note, likely correctly, that any associated administrative costs will ultimately be passed along to consumers as higher premiums or, in the case of self-insured plans, higher plan administration fees. (For self-insured plans, the plan administrator, rather than the employer itself, generally handles APCD submission).

Our conversations with stakeholders indicated that submitting to an APCD involves three main steps. First, the insurer or plan administrator must extract the required information from its data systems and put that information in the format required by the APCD. Historically, each state's APCD has set its own data submission format, although in recent years the APCD Council has worked with state APCDs to develop a Common Data Layout that may be used in more states in years to come.²³ Second, the submitting entity must actually transmit those data to the APCD, a process that generally uses standard technical protocols. Finally, the submitting entity must respond to any post-submission quality control inquiries. Our stakeholder conversations indicated that it is common for a payer to be required to submit data multiple times, either because the payer's data fails automated data quality checks at the time of submission or because the APCD's post-submission quality control checks identify issues that require resubmission.

¹⁸ Center For Health Information And Analysis, "Enrollment in Health Insurance," <https://www.chiamass.gov/enrollment-in-health-insurance/> (last visited September 17, 2020).

¹⁹ Sarah H. Gordon, Benjamin D. Sommers, Ira Wilson, Omar Galarraga, and Amal N. Trivedi, "The Impact of Medicaid Expansion on Continuous Enrollment: a Two-State Analysis," 34 JOURNAL OF GENERAL INTERNAL MEDICINE 1919 (June 21, 2019), <https://link.springer.com/article/10.1007/s11606-019-05101-8>.

²⁰ Almost all states maintain these databases for inpatient stays, and most states also have such databases for emergency department visits and hospital-based outpatient surgeries. See, "HCUP Fact Sheet," *Healthcare Cost and Utilization Project*, https://www.hcup-us.ahrq.gov/news/exhibit_booth/HCUPFactSheet.pdf (last visited September 17, 2020).

²¹ "Achieving States' Goals for All-Payer Claims Databases," *Anthem Public Policy Institute*, June 2018, https://www.antheminc.com/cs/groups/wellpoint/documents/wlp_assets/d19n/mzq1/~edisp/pw_g345393.pdf.

²² The ERISA Industry Committee, "Comments on the "Lower Health Care Costs Act of 2019," June 5, 2019, <https://www.eric.org/uploads/doc/resources/06-05-19%20ERIC%20Comments%20on%20HELP%20Draft%20Final.pdf>.

²³ All-Payer Claims Database, "Common Data Layout," <https://www.apcdouncil.org/common-data-layout> (last visited September 17, 2020).

We are unaware of any estimates of the cost of submitting data to an APCD. We note, however, that the activities involved in APCD submission are relatively routine for insurers and plan administrators. Data like these are shared with various vendors associated with day-to-day operation of health plans. Notably, many insurers and plan administrators also submit essentially identical data to commercial or non-profit databases (e.g, IBM MarketScan, Blue Health Intelligence or the Health Care Cost Institute). The incremental cost in submitting to an APCD may, therefore, be modest. In any case, we discuss steps federal policymakers could take to minimize these costs below.

Upward Pressure on Negotiated Prices

Another commonly expressed concern is that APCDs may result in public disclosure of the prices negotiated between health care providers and health insurers and that this disclosure may put upward pressure on negotiated prices.²⁴ That upward pressure could arise in two ways.²⁵ First, price disclosure may facilitate tacit collusion by providers by making it harder for a provider that lowers its price to hide that fact from its competitors. Second, in some cases, price disclosure could change the landscape of provider-insurer negotiations in ways that increase prices; for example, providers could become more reluctant to accept low prices from any given insurer for fear that other insurers will see that they are willing to accept low prices and demand low prices too, or a low-priced provider could learn that an insurer is willing to pay other providers higher prices and demand similarly high prices.

On the other hand, there are also ways that making price information more broadly available could reduce negotiated prices. First, as noted above, some states use APCD data to drive consumer-facing tools designed to enable patients to seek out lower-priced providers. If consumers did shift to lower-priced providers, this could lower prices directly and could put pressure on providers to reduce prices. Second, in some cases, price disclosure could change the landscape of provider-insurer negotiations in ways that reduce prices, rather than increase them as discussed in the last paragraph; for example, insurers could become less willing to pay any given provider high prices for fear that other providers will see that they are willing to pay high prices and demand high prices too, or an insurer that currently pays a provider a high price could learn that the provider accepts lower prices from other insurers and demand a similarly low price.

Some recent empirical research has suggested that greater price transparency may, on net, cause small reductions in prices, but the effects of price disclosure may be context-dependent, and this question is far from settled.²⁶ In any case, if federal policymakers are concerned that price disclosure may have downsides, they could take steps to limit APCDs' ability to disclose provider-level price information, although this would inhibit some uses of APCD data.

²⁴ "Achieving States' Goals for All-Payer Claims Databases," *Anthem Public Policy Institute*, June 2018, https://www.antheminc.com/cs/groups/wellpoint/documents/wlp_assets/d19n/mzq1/~edisp/pw_g345393.pdf; Chamber of Commerce of the United States of America, Letter to Senators Alexander and Murray regarding the Lower Health Care Costs Act, June 17, 2019, https://www.uschamber.com/sites/default/files/190617_lowerhealthcarecostsact_senatehelpcommittee.pdf.

²⁵ For an overview of the economic logic and evidence behind these arguments, see "Amendments to the Minnesota Government Data Practices Act Regarding Health Care Contract Data," *Office of Policy Planning, Bureau of Competition, Bureau of Economics*, June 29, 2015, https://www.ftc.gov/system/files/documents/advocacy_documents/ftc-staff-comment-regarding-amendments-minnesota-government-data-practices-act-regarding-health-care/150702minnhealthcare.pdf.

²⁶ Christopher M. Whaley, "Provider Responses to Online Price Transparency," 66 *JOURNAL OF HEALTH ECONOMICS* 241 (July 2019), <https://www.sciencedirect.com/science/article/abs/pii/S0167629618310476>; Zach Y. Brown, "Equilibrium Effects of Health Care Price Information," 101 *REVIEW OF ECONOMICS AND STATISTICS* 699 (October 2019), https://www.mitpressjournals.org/doi/full/10.1162/rest_a_00765?casa_token=OlbiZgME8_gAAAAA%3AWwnoEMMIIsD_7YhOPSIJ5otMLz31TUuQONvUiHhq69VWIet8GqIF5Q7YYNmXFjxxlzdR2z_xrOEbsq.

Privacy and Security

Stakeholders also frequently raise concerns about the privacy of claims information submitted to an APCD,²⁷ fearing disclosure of individual-level information, such as through a computer systems breach or as a result of researcher misconduct.²⁸ While this is indeed a theoretical possibility, these concerns can be addressed through appropriate data privacy security safeguards. Indeed, we are unaware of any significant claims data breaches in the Medicare and Medicaid programs, both of which hold similarly sensitive claims records and use them for similar purposes. We discuss how federal policymakers might approach privacy and security protection in the context of a national APCD later in this paper.

Data Uses

Stakeholders also sometimes raise concerns about the fact that APCDs generally make data broadly available for public reporting, research, and policymaking, rather than restricting use of the data to a narrow set of pre-specified “use cases.” This open-ended mandate makes some stakeholders – particularly those whose activities are subject to more careful scrutiny through an APCD – uncomfortable. However, much of the value of an APCD arises from its ability to offer a flexible tool to support a broad array of efforts to better understand and ultimately improve the health care system. Because the health care system is complex and because circumstances change over time, any attempt to pre-specify the full set of potential applications of APCDs would likely leave out many high-value applications of APCD data.

A related concern, while not always made explicit, is that a government entity ought not have access to this type of health care data, particularly in the context of the broad mandate envisioned for APCDs. There is little that policymakers can do to mitigate this concern while still realizing the potential of APCDs. We do note, however, that even before the advent of state APCDs, state governments held claims data for their Medicaid programs and the federal government held claims data for the Medicare program, so it is routine for government entities to collect and hold claims data, albeit generally for the purposes of direct program administration.

Unstated Stakeholder Objections

While the preceding objections to APCDs are the ones most often raised publicly, some stakeholders’ views of APCDs may also reflect more parochial concerns. First, some health care providers may worry that broader availability of data on the prices of health care services may engender support for policies to reduce those prices. Second, as noted above, many insurers currently sell their claims information or data products derived from that claims information to third parties. For example, Blue Cross Blue Shield plans do so through their Blue Health Intelligence affiliate, while United Healthcare does the same through its Optum subsidiary, and many plans also sell data to data warehouses like IBM MarketScan that then resell those data to other entities. Plans may worry that broader availability of data through APCDs would reduce the prices they can demand. Third, some dominant insurers may worry that if information on prices and utilization in their markets become more widely available, other insurers could use that information to enter those markets or otherwise compete more effectively. In general, there is not a strong policy rationale for changing APCD policy to address parochial concerns like these, but policymakers should be aware that stakeholders may harbor them.

²⁷ Federation of American Hospitals, Letter to Senators Alexander and Murray Regarding the Lower Health Care Costs Act, June 5, 2019, https://www.fah.org/fah-ee2-uploads/website/documents/FAH_Response_-_HELP_Health_Care_Cost_Reduction_Discussion_Draft_%28FINAL%29.pdf; “Achieving States’ Goals for All-Payer Claims Databases,” *Anthem Public Policy Institute*, June 2018, https://www.antheminc.com/cs/groups/wellpoint/documents/wlp_assets/d19n/mzq1/~edisp/pw_g345393.pdf; Chamber of Commerce of the United States of America, Letter to Senators Alexander and Murray regarding the Lower Health Care Costs Act, June 17, 2019, https://www.uschamber.com/sites/default/files/190617_lowerhealthcarecostsact_senatehelpcommittee.pdf.

²⁸ Peter Swire, “Possible Privacy, Cybersecurity, and Data Breach Issues in the Proposed National Medical Claims Database Under Section 303 of S. 1895,” September 27, 2019, <https://peterswire.net/wp-content/uploads/Swire-white-paper.S-1895-privacy-security.2019.pdf>.

Limitations of Existing APCDs

Existing APCDs generate real benefits but also have limitations that reduce their utility for public reporting, research, and policymaking. One limitation—which has received increasing attention from federal policymakers—stems from the Supreme Court’s 2016 decision in *Gobeille v. Liberty Mutual Insurance Co.*, which held that states could not require data collection from non-governmental self-insured group health plans.

But existing APCDs also have other limitations that predate the *Gobeille* decision and arise from the fact that current APCDs are state entities that operate almost entirely independently of one another. These limitations, particularly the fact that it is challenging to combine data from multiple APCDs, reduce APCDs’ utility in many applications, especially at the federal level, while also increasing administrative costs. This section discusses each set of limitations in turn.

The *Gobeille* Decision and its Consequences

Gobeille was a 6-2 decision, with Justices Roberts, Kennedy, Thomas, Breyer, Alito, and Kagan in the majority. This section considers the Court’s holding and the decision’s impact on APCDs.

The Court’s Decision in *Gobeille v. Liberty Mutual*

The federal Employee Retirement Income Security Act of 1974 (ERISA) regulates benefits offered by employers, including employer health plans (also called group health plans). ERISA includes a famously broad preemption clause that bars states from implementing any laws that “relate to” employee benefits.²⁹ However, ERISA also specifies that state laws that “regulate[] insurance” are not preempted by federal law,³⁰ leading to an uneasy arrangement where states have jurisdiction over the sale of insured health benefits to employers, but not over the employer’s group health plan itself. As a result, states can generally use their authority to regulate insurance to control *insured* employer health benefits, but state law cannot reach self-insured group health plans. And self-insurance is common: 61% of people with health coverage from an employer were in a self-insured plan in 2019.³¹

While ERISA does place limits on how states interact with self-insured group health plans, the scope of ERISA preemption was unclear. For a number of years, states took the position that APCD data collection did not violate ERISA. As a practical matter, state laws generally placed the reporting obligation on the entity that administered benefits and paid claims, which for almost all self-insured group health plans is a third-party (generally an insurance company), not on the plan itself. Thus, states could maintain that the data collection laws were not “related to” a group health plan as the Supreme Court has defined that concept in jurisprudence dating back to the 1990s.³²

However, in 2011, a self-insured plan directed its administrator not to submit data to Vermont, the administrator was subpoenaed by the state, and the self-insured plan then sued, claiming that the requirement to submit to the state’s APCD was preempted by ERISA. The Second Circuit agreed.³³ The case reached the Supreme Court, and in March 2016 the Supreme Court held that ERISA preempted state laws that required data submission associated with a self-insured plan.³⁴

²⁹ 29 U.S.C. § 1144(a).

³⁰ 29 U.S.C. § 1144(b)(2).

³¹ “2019 Employer Health Benefits Survey,” *Kaiser Family Foundation*, September 25, 2019, <https://www.kff.org/report-section/ehbs-2019-section-10-plan-funding/>.

³² Given the Court’s prior decisions, this argument was more plausible than it may initially seem. The Court has criticized “uncritical literalism” in applying the phrase “relate to,” and demanded a more contextual inquiry into whether a state law duplicates ERISA requirements in assessing preemption. See, e.g., Brief for Petitioner, *Gobeille v. Liberty Mutual Insurance Co.*, <https://www.scotusblog.com/wp-content/uploads/2015/09/14-181ts.pdf>.

³³ *Liberty Mutual Insurance Co. v. Donegan*, 746 F.3d 497 (2d Cir. 2014).

³⁴ *Gobeille v. Liberty Mutual Insurance Co.*, 577 U.S. ____ (2016)

The Court’s majority opinion concluded that state laws requiring administrators to submit data about self-insured plans were “related to” a group health plan within the meaning of ERISA. Under the Court’s prior ERISA jurisprudence, one important question was the extent to which the state law overlapped with “the objectives of the ERISA statute.” The Court emphasized that ERISA included record-keeping and reporting requirements and that while current ERISA regulations did not directly compel self-insured plans to submit claims data to federal regulators, the statute provided authority for the federal government to require submission of additional data.³⁵ Given this overlap between the “objectives” of ERISA and these state laws, the states’ reporting requirements were preempted.

The scope of federal data collection authority was also addressed in a separate concurrence by Justice Breyer. Breyer noted that federal law allowed the government to collect information related to health care claims that was very similar to the data state APCDs were collecting, and he wrote to emphasize that this federal authority provided a pathway for states to access the data they sought. He noted that the federal government could collect claims data and conduct analysis on behalf of the states, share data with the states, or “delegate” authority to collect data to “a particular state.”³⁶

Consequences of the Gobeille Decision

Today, state APCDs are continuing to collect and analyze data from insured group health plans, which represent about 39% of the employer market nationally.³⁷ State APCDs can also require data collection from self-insured non-federal governmental health plans (i.e., state and local government employee health plans), which are not regulated under ERISA or affected by ERISA preemption. These plans account for a significant fraction of total enrollment in most states since state government and public universities are major employers almost everywhere in the country. Finally, states generally allow voluntary submission from self-insured plans (other than governmental plans).

We are unaware of any comprehensive data on how many self-insured plans still submit to APCDs, but fragmentary evidence suggests that states are receiving a limited amount of data. For example, Maryland reports that it collects data for 25-30% of self-insured enrollees, primarily from governmental plans, and Massachusetts estimates that it collects data for about 25% of self-insured enrollees.³⁸ Our conversations with stakeholders suggest that some large employers, particularly those that have an institutional connection to the APCD mission like health systems or universities, do opt in to data-sharing with state APCDs and that some states have also worked with Chambers of Commerce or other local stakeholders to encourage employers to opt in. However, stakeholders suggest that these efforts have borne limited fruit, and data collection from non-governmental self-insured plans is fairly limited and non-representative.

Nor are there obvious steps states could take to substantially increase submission by self-insured plans. States could try to require third-party administrators to submit self-insured plan data unless the employer affirmatively opts out of submission, but given that many plan administrators may prefer not to submit data for the various reasons discussed above, opt outs would likely be common. Alternatively, states could require that any insurance company that wished to sell *insured* health benefits in the state must incorporate submission to the state APCD into its contracts with employers for which it administers self-insured plans. But this strategy would not reach all self-insured plans, and courts might conclude that this type of regulatory bank shot was also preempted under ERISA.

³⁵ Id. (“The State’s law and regulation govern plan reporting, disclosure, and—by necessary implication—recordkeeping. These matters are fundamental components of ERISA’s regulation of plan administration.”)

³⁶ *Gobeille v. Liberty Mutual Insurance Co.*, 577 U.S. ____ (2016) (Breyer, J., concurring).

³⁷ “2019 Employer Health Benefits Survey,” *Kaiser Family Foundation*, September 25, 2019, <https://www.kff.org/report-section/ehbs-2019-section-10-plan-funding/>.

³⁸ The Maryland Health Care Commission, “MCDB Data Release,” January 14, 2020, https://mhcc.maryland.gov/mhcc/pages/apcd/apcd_data_release/apcd_data_release_mcdb.aspx; “Massachusetts All-Payer Claims Database, 2014-2018 Documentation Guide,” *Center for Health Information and Analysis*, February 2020, <https://www.chiamass.gov/assets/docs/p/apcd/apcd-8.o/APCD-Release-8-Documentation-Guide.pdf>.

The loss of self-insured data makes state APCDs less useful than they were prior to *Gobeille* in at least four important respects:

- **Loss of comprehensiveness:** APCDs that lack complete self-insured data are not true all-payer databases since, as noted earlier, about one-third of all covered people are enrolled in self-insured group health plans. That large gap makes it difficult or impossible to use the APCD to track aggregate insurance enrollment or to study transitions among different coverage types, eliminating one of the major advantages of APCDs relative to other data sources.
- **Loss of representativeness:** The enrollees omitted when APCDs lose access to self-insured data are a non-random subset of people with employer coverage. Notably, large employers are far more likely to be self-insured; in 2019, just 17% of health insurance enrollment was in self-insured plans among firms with fewer than 200 workers, compared to 80% for firms with 200 or more workers.³⁹ The share of workers in self-insured plans also varies widely by industry and region, presumably in part reflecting variation in the firm size distribution across those categories. Consequently, state APCDs that lack self-insured data may offer a skewed picture of the market as a whole and will generally be unsuitable for studying how claims spending varies across firms of different types, particularly larger and smaller firms.
- **Smaller sample sizes:** Due to the large number of covered lives in self-insured plans, exclusion of self-insured data from state APCDs also substantially reduces the size of those databases. Smaller sample sizes can create problems for many analyses but are particularly problematic for efforts to measure the efficiency or quality of care at the provider level since the total number of patients seen by many providers (across all plan types) is often modest.
- **Inability to study differences between insured and self-insured plans:** APCDs that lack data from self-insured plans cannot be used to compare insured and self-insured plans. Because of ERISA preemption, insured health plans are subject to many state regulations that self-insured plans are not, so comparing outcomes under the two types of plans can provide useful insights about the effects of those regulations.⁴⁰ Some research has also found that health insurance companies behave differently when acting as third-party administrators for self-insured plans than when they sell insured coverage, a pattern that merits greater study.⁴¹

Limitations of Relying on a Patchwork of State APCDs

The *Gobeille* decision was a significant blow to state APCDs, but even the APCDs that existed before the decision fell short of realizing APCDs' full potential. Because existing APCDs are run by the states, they operate independently of one another and do not exist everywhere. In several ways, this state of affairs has made existing APCDs less effective than they could be in supporting public reporting, research, and policymaking, while increasing administrative costs for both payers and governments.

Challenges in Combining Data from Multiple States

Many potential applications of APCD data require combining data for multiple states. Most prominently, research aimed at learning about the effects of state policies commonly involves comparing outcomes in states that have implemented a particular policy to states that have not and thus requires data for multiple states. In other cases, data from multiple states may be necessary to achieve adequate statistical power, particularly when studying phenomena that operate at the provider

³⁹ "2019 Employer Health Benefits Survey," *Kaiser Family Foundation*, September 25, 2019, <https://www.kff.org/report-section/ehbs-2019-section-10-plan-funding/>.

⁴⁰ See, e.g., Colleen L. Barry, Andrew J. Epstein, Steven C. Marcus, Alene Kennedy-Hendricks, Molly K. Candon, Ming Xie, and David S. Mandell, "Effects Of State Insurance Mandates On Health Care Use And Spending For Autism Spectrum Disorder," 36 *HEALTH AFFAIRS* 1754 (October 2017), <https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2017.0515>.

⁴¹ Stuart V. Craig, Keith Marzilli Ericson, and Amanda Starc, "How Important Is Price Variation Between Health Insurers?," *The National Bureau of Economic Research*, October 2018, <https://www.nber.org/papers/w25190>.

or geographic market level (since, in these cases, the relevant metric of sample size is the number of providers or markets represented in the data set). Similarly, communities interested in understanding in broad terms how their health care systems compare to others will often benefit greatly from being able to compare to communities in other states. Even studies with a purely local focus will sometimes require data from multiple state APCDs if a metropolitan area of interest crosses state boundaries.

Unfortunately, despite the benefits of combining data from multiple state APCDs, we are aware of few studies that have done so (and even those studies that do combine data from multiple state APCDs generally only use data from a small number of APCDs).⁴² The dearth of such studies likely reflects two main barriers. First, and likely most important, while most state APCDs allow researchers to access their data if certain conditions are met, as noted earlier, each state has its own application process, its own restrictions on how data can be used, and its own fees for data access. Consequently, accessing multiple states' databases may require a substantial investment of both time and funds. Second, each state APCD collects and stores data in slightly different ways. Thus, adding an additional state to a research project generally requires substantial additional researcher effort to account for the idiosyncratic features of that particular state's data. For both these reasons, research projects that use more than one state database are likely to be prohibitively difficult in most cases.

Incomplete Geographic Coverage

Currently, 27 states lack APCDs, and half the country's population lives in a state without an APCD. The most direct consequence of the fact that some states lack APCDs is that these states cannot use APCD data to support research and policy efforts aimed at improving their own health care systems.

But the fact that many states lack APCDs also impedes national efforts to improve the health care system. For example, researchers have used APCD data to evaluate state-level policies, which can offer lessons that are useful to other states and to federal policymakers.⁴³ APCD data cannot support such studies in states that lack them. And even where barriers to combining data from multiple states can be overcome, the lack of truly national data constrains the sample sizes available to researchers.

Incomplete APCD coverage also limits the utility of APCDs for federal policymaking, even if the data access concerns considered below can be overcome. For example, as noted earlier, some states have used their APCDs to set out-of-network payment standards in legislation addressing surprise billing. But because APCDs do not exist everywhere, that option is not available to federal policymakers. Instead, they have ended up pursuing other approaches to setting payment standards, such as having each insurer compute a standard based on its own data, which have important downsides.⁴⁴

Inaccessibility to Federal Policymakers

APCD data may be of use to federal policymakers in areas well beyond surprise billing. The comprehensive picture they provide of the commercial health insurance market can help inform policy analysis and policy development work by both the executive branch and legislative agencies like the

⁴² For a couple of notable exceptions, see Sarah Gordon, Benjamin Sommers, Ira Wilson, Omar Galarraga, and Amal N. Trivedi, "The Impact of Medicaid Expansion on Continuous Enrollment: a Two-State Analysis," 34 *JOURNAL OF GENERAL INTERNAL MEDICINE* 1919 (June 21, 2019), <https://link.springer.com/article/10.1007/s11606-019-05101-8>; Maria de Jesus Diaz-Perez, Rita Hanover, Emilie Sites, Doug Rupp, Jim Courtemanche, and Emily Levi, "Producing Comparable Cost and Quality Results From All-Payer Claims Databases," 25 *AMERICAN JOURNAL OF MANAGED CARE* 138 (May 2, 2019), <https://www.ajmc.com/journals/issue/2019/2019-vol25-n5/producing-comparable-cost-and-quality-results-from-allpayer-claims-databases>.

⁴³ See, e.g., Keith Marzilli Ericson, Amanda Starc, "How Product Standardization Affects Choice: Evidence From The Massachusetts Health Insurance Exchange," 50 *JOURNAL OF HEALTH ECONOMICS* 71, December 2016, <https://www.sciencedirect.com/science/article/abs/pii/S0167629616302156>.

⁴⁴ Loren Adler, Matthew Fiedler, Paul B. Ginsburg, and Christen Linke Young, "Comments on the Lower Health Care Costs Act of 2019," *Brookings Institution*, June 6, 2019, <https://www.brookings.edu/opinions/comments-on-the-lower-health-care-costs-act-of-2019/>; Loren Adler, Matthew Fiedler, Paul B. Ginsburg, and Christen Linke Young, "Comments on the No Surprises Act," *Brookings Institution*, May 29, 2019, <https://www.brookings.edu/opinions/comments-on-the-no-surprise-act/>.

Congressional Budget Office. APCD data may also be useful for various “operational” purposes, including anti-trust enforcement by the Department of Justice and Federal Trade Commission.

Unfortunately, because existing APCDs are controlled by the states, they are not routinely available to federal policymakers. To our knowledge, no state provides a specific process by which a federal government agency can gain access to APCD data, although a federal agency might be able to access data through the process available to researchers. Even where processes do exist, federal agencies will face the same challenges researchers face in trying to combine data from multiple states, although their greater resources may increase their ability to overcome them.

Inability to Exploit Economies of Scale

The existence of multiple state APCDs also necessitates some administrative duplication. Each state must develop its own data submission policies and protocols, build and maintain its own data systems, manage and clean its own data, as well as produce its own publications and statistical reports based on the data collected.⁴⁵ For their part, payers that operate in multiple states incur higher costs to conform to each state’s data submission requirements and respond separately to post-submission inquiries regarding data quality issues. The development of the APCD Council’s Common Data Layout (CDL), which was described above, may reduce duplicative effort in some areas, but will not in others, and it remains to be seen how widely the CDL will ultimately be adopted.⁴⁶

Because APCD budgets are relatively modest and there is reason to believe that payers’ submission costs are modest too, the aggregate cost of this duplication may be small, at least relative to the potential benefits of APCDs and health care spending. Rather, the more important way that state APCDs’ limited scale negatively affects their work may be by preventing them from making certain types of investments. For example, as noted earlier, APCD data can be used to produce reports on aggregate trends in health care spending, prices, utilization, and quality across different service types or geographic areas, which can be valuable to a broad array of users, ranging from researchers and policymakers to employers and health insurers. Investing in the staff to produce these types of reports may sometimes be challenging for states, particularly small states. But because producing these types of reports for many geographic areas is only modestly more resource intensive than producing them for a single geographic area, these types of investments would likely be more feasible for an APCD with broader geographic scope.⁴⁷ For similar reasons, APCDs with broader geographic scope may be able to justify investing more in efforts to improve data quality since those efforts could be leveraged across a larger database. States’ limited financing capacity may also lead APCDs to rely too heavily on fees from data users, which may reduce the data’s accessibility for research purposes.

Federal Policy Options

Recent years have seen bipartisan federal interest in improving APCDs (see Box 2), which suggests that changes in federal policy toward APCDs are indeed possible. In this section, we consider steps federal policymakers could take to make progress in this area. We consider three broad approaches.

The first approach would, through either legislative or administrative action, restore state APCDs’ ability to require submission of data for self-insured plans, allowing state APCDs to once again provide a comprehensive picture of how health care services are received and paid for in their states. However, this approach would not address the various other limitations of the existing APCDs discussed above or would address these other limitations in an incomplete or cumbersome way.

⁴⁵ In practice, some vendors serve multiple state APCDs, which may reduce duplication to some degree.

⁴⁶ All-Payer Claims Database, “Common Data Layout,” <https://www.apcdouncil.org/common-data-layout> (last visited September 17, 2020).

⁴⁷ Notably, the Health Care Cost Institute has produced reports with a national scope.

Box 2: Recent Federal Proposals Related to APCDs

Federal policymakers of both parties have shown interest in making APCDs more effective. Most of those proposals have been primarily focused on addressing the problems created by the *Gobeille* decision, but some would also have helped address other problems with current APCDs.

In July 2016, just three months after the *Gobeille* decision, the Obama Administration promulgated a proposed regulation updating its data collection standards for all employee benefits. It did not propose specific language related to the collection of APCD-like information but did seek public comments on what changes it should make to health plan reporting “in light of the Supreme Court’s recent decision in *Gobeille v. Liberty Mutual Insurance Co.*”⁴⁸ Commenters, including those representing both self-insured plans and APCDs, took this as an indication that the Department of Labor was considering policy along the lines indicated in Justice Breyer’s concurrence that would allow state APCDs to regain access to self-insured plan data.⁴⁹ The federal government subsequently delayed the update process and withdrew this rule in the fall of 2019,⁵⁰ seemingly due to general concerns about regulatory burden not specific to APCDs.

A more concrete proposal was introduced in the summer of 2019 by Senators Lamar Alexander (R-TN) and Patty Murry (D-WA), leaders of the Senate Health, Education, Labor, and Pensions (HELP) committee.¹ In their Lower Health Care Costs Act, they proposed creation of what would essentially be a national APCD, containing data from all states and housed in a non-profit entity. Under their proposal, self-insured plans would be required to submit data to this system. These data would be available to state APCDs on the condition that they provide state-level data on Medicaid and insured plans. States could also require insured plans and other payers to submit data directly. The proposal established a process for research use and provided \$15 million per year in funding.⁵¹

In addition, in early 2020, Representative Dan Lipinski (D-IL) proposed legislation that focused narrowly on removing the barriers to APCDs created by the *Gobeille* decision. His bill would have simply modified ERISA to allow state APCDs to resume collecting data from self-insured plans.⁵² This legislation also included federal grants to support new and existing state APCDs.

Two additional approaches aim to address both the problems created by the *Gobeille* decision and the limitations of relying on an uncoordinated patchwork of state APCDs that pre-dated *Gobeille*. Policymakers could build a national APCD, which could fully address all of the limitations of existing APCDs discussed above. Alternatively, policymakers could seek to harmonize existing state APCDs and make state APCD data available through a federal clearinghouse, which would address some, but not all, of these limitations.

Table 1 summarizes these three potential changes to federal policy toward APCDs, as well as two other prominent proposals. The remainder of this section discusses these proposals in much greater detail.

⁴⁸ Employee Benefits Security Administration, Internal Revenue Service, Pension Benefit Guaranty Corporation, “Proposed Revision of Annual Information Return/Reports,” 81 Fed. Reg. 47533 (July 21, 2016), <https://www.federalregister.gov/documents/2016/07/21/2016-14893/proposed-revision-of-annual-informationreturnreports>.

⁴⁹ See, e.g., APCD Council, Comment Letter Regarding Employee Benefits Security Administration, Annual Reporting and Disclosure Proposed Rule, October 12, 2016, <https://www.regulations.gov/document?D=EBSA-2016-0010-0046>; BlueCross BlueShield Association, Comment Letter Regarding Employee Benefits Security Administration, Annual Reporting and Disclosure Proposed Rule, December 12, 2016, <https://www.regulations.gov/document?D=EBSA-2016-0010-0130>.

⁵⁰ 29 C.F.R. § 2520, Fall 2019 Unified Agenda, “Revision of the Form 5500 Series and Implementing Related Regulations Under the Employee Retirement Income Security Act of 1974,” <https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=201910&RIN=1210-AB63>.

⁵¹ The bill also authorized grants to states to establish and maintain APCDs but did not appropriate funds for this purpose.

Table I: Comparison of Potential Federal Policy Changes to Improve APCDs

	Administrative or legislative?	Allows state APCDs to collect self-insured data?	Pre-Gobeille Limitations of APCDs			
			Facilitates combining data from many states?	Expands APCD coverage?	Ensures federal gov't can access APCD data?	Benefits from economies of scale?
Policy Approaches Considered in This Paper						
Enable state collection of self-insured data	Either	Yes, directly	No (or limited)	No	No (or limited)	No (or limited)
Build a national APCD in a federal agency or non-profit	Either, but legislative path more likely	Yes, from national APCD	Yes	Yes, fully national	Yes	Yes
Harmonize state APCDs and create a federal clearinghouse	Either, but legislative path more likely	Yes, directly	Yes, with some limitations	Possibly, via grant funding	Yes	No (or limited)
Other Approaches						
Collect self-insured data nationally and share with state APCDs (Senate HELP proposal)	Either, but legislative path more likely and needed for grants	Yes, from national APCD	Yes, for self-insured data, other data with limitations	Yes, for self-insured data, but not for other data	Yes	Partially
“Federated” approach with payers holding their own data and executing research queries	Legislative	Only through research queries	No	No	No	No

Enable State Collection of Self-Insured Data

Policymakers that wish to restore something like the pre-*Gobeille* status quo could pursue one of two broad paths. The simplest approach would be for Congress to amend ERISA to restore states' ability to obtain data for self-insured plans, but agencies could achieve a similar outcome through rulemaking.

Legislative De-Preemption

The Supreme Court's decision in *Gobeille* was a statutory decision (that is, it was based on the Court's interpretation of the text of ERISA), so Congress could restore states' ability to require data submission from self-insured plans by simply amending ERISA. Specifically, the law's preemption clause could be modified to specify that state laws requiring data submission to an APCD are not preempted; Representative Lipinski (D-IL) introduced legislation doing exactly that in early 2020.⁵²

Congress has changed the scope of ERISA preemption in a similar past instance: in 1981, the Supreme Court held that Hawaii's employer mandate as preempted by ERISA,⁵³ and in 1983, Congress modified ERISA's preemption clause to declare that the state law was exempt from preemption.⁵⁴ Further, such an approach would not represent a radical departure from how ERISA's preemption clause has functioned historically. Since the 1990s, ERISA's preemptive scope has been fairly limited in health policy because it is an area of traditional state regulation,⁵⁵ so lawmakers need not worry that modifying the preemption clause in this way would disrupt the uniformity of the statutory scheme.

Congress could allow state APCDs to resume data collection with no restrictions. Alternatively, Congress could define the scope of de-preemption more narrowly, allowing states to avoid preemption only when state law meets certain criteria. In principle, this approach could help to address some of the limitations of state APCDs that predated *Gobeille* or ameliorate some stakeholder concerns. This would be a more significant departure from how ERISA's preemption clause has historically been drafted, but it is feasible.⁵⁶

Congress may wish to consider this type of approach in two main areas:

- **Data format and submission processes:** To address concerns about the burden of data collection, Congress could condition de-preemption on states collecting data in a fairly standardized format. It would be unwise to legislate use of any specific format, such as the APCD Council's Common Data Layout, since that would preclude changes as technology improved. Rather, states could be required to collect data in a format established by federal agencies (likely the Department of Labor, in consultation with the Department of Health and Human Services) through regulation. This standardized format might naturally be the CDL initially but could evolve over time. The standardized format could also, in principle, make some allowance for states to collect state-specific data elements (as the CDL does).

⁵² Transparency and Accountability in Health Care Costs and Prices Act of 2020, H.R. 6004, 116th Congress (2020), <https://www.congress.gov/bill/116th-congress/house-bill/6004>.

⁵³ *Agsalud v. Standard Oil Co.*, 454 U.S. 801 (1981).

⁵⁴ 29 U.S.C. § 1144(b)(5).

⁵⁵ See, e.g., *Travelers*.

⁵⁶ Note that Justice Thomas has articulated Commerce Clause concerns about ERISA's preemption clause as a whole. For example, in his concurrence in *Gobeille*, he wonders about the extent to which "Congress can exempt ERISA plans from state regulations that have nothing to do with interstate commerce," and he has called upon the Court to reconsider its ERISA jurisprudence as a whole. *Gobeille v. Liberty Mut. Insurance Co.*, 577 U.S. ____ (2016) (Thomas, J., concurring). Legislation that conditions de-preemption on a variety of technical considerations may further exacerbate these concerns, though Justice Thomas's views do not appear to command a majority.

Similarly, Congress could also consider measures to standardize the process by which states submit data to APCDs. This could include standardizing the schedule on which payers are required to submit data or how payers transmit data, although the benefits of standardizing the latter may be small.⁵⁷ As with data submission formats, it would be unwise to legislate a particular schedule or transmission process, but Congress could direct the agencies to monitor specific aspects of the data submission process and provide authority for them to standardize elements of that process if opportunities to streamline it became apparent.

Congress could also consider limiting de-preemption to data collection from administrators that are either responsible for a minimum number of self-insured lives in the state or have any insured business in the state. For other administrators, the benefits of collecting the additional data may be small relative to the additional administrative burden created. However, it would be important that any exclusion threshold be set at a reasonably low level in order to ensure that APCDs remain representative of the state market.

- **Facilitating data use:** Congress could also condition de-preemption on APCDs having a suitable process for making data available to federal policymakers, researchers, and potentially other data users to access data. For example, states could be required to provide access to federal policymakers, adopt a harmonized application process for researchers that makes it easier for researchers to obtain data from multiple states, or produce a minimum set of public reports with aggregate data on health care spending. As above, it would likely be unwise to specify the details of these requirements in legislation, so Congress may wish to grant general authority to federal agencies to establish requirements like these in regulation.

Implementing these types of standards would likely reduce the burden on submitting entities and might make it somewhat easier for policymakers and researchers to access data, although they would likely accomplish much less in either area than proposals that would create a national APCD infrastructure. On the other hand, conditions of this kind, particularly conditions that would mandate that states offer broader data access, may require changes in some state laws since not all states provide research use on the same terms. That could meaningfully delay or even block access to self-insured data by some state APCDs. Changes to data collection formats and processes would also involve some transition costs for states and for payers. Efforts by the agencies to define and oversee these standards would also carry some opportunity costs and distract agency staff from other priorities. All of these costs should be weighed against the benefits of greater standardization going forward.

Agency De-Preemption

In the absence of new federal legislation, federal agencies have authority to facilitate state collection of data from self-insured plans through regulation. The Supreme Court's majority opinion in *Gobeille* and Justice Breyer's concurrence both highlighted the federal government's authority under ERISA to establish reporting requirements for all employee benefits. The majority opinion noted that ERISA allows the Department of Labor to compel reporting of "such data or information [that] is necessary to carry out the purposes of" ERISA⁵⁸ and to use these data "for statistical and research purposes, and [to] compile and publish such studies, analyses, reports, and surveys."⁵⁹ The Court also noted that the Affordable Care Act included new reporting requirements for group health plans that could encompass data related to health care claims.⁶⁰ This is perhaps a more expansive view of the ERISA data collection authorities than the Department of Labor has previously adopted, but the majority opinion reflects the Court's view that ERISA grants the agency authority to collect granular information from group health plans.

⁵⁷ States generally use routine technical transactions for submission and whatever burden exists arises from the mechanics of establishing a connection to the APCD, which should generally be relatively easy.

⁵⁸ 29 U.S.C. § 1024(a)(2)(B).

⁵⁹ 29 U.S.C. § 1026(a).

⁶⁰ 29 U. S. C. § 1185d; 42 U. S. C. § 300gg-15a

Justice Breyer's concurrence addresses these issues even more directly. He explicitly says that the authorities cited by the majority allow the Department of Labor to require self-insured plans to report data that mirrors the data collected by state APCDs. Further, he crafts what he views as a plausible path for state APCDs to continue to access self-insured data, mediated by the Department of Labor. This is consistent with a theme that Breyer has articulated in a series of concurrences beginning in the mid-1990s: that federal agencies can play an important role in helping courts to understand the preemptive scope of statutes under their jurisdiction.⁶¹

Specifically, Breyer's concurrence envisions that federal government could collect APCD-like information and share the data with states, as appropriate, or the federal government could craft a path for states to access the data directly by "delegating" authority to the states:

I see no reason why the Secretary of Labor could not develop reporting requirements that satisfy the States' needs, including some State-specific requirements, as appropriate. Nor do I see why the Department could not delegate to a particular State the authority to obtain data related to that State, while also providing the data to the Federal Secretary for use by other States or at the federal level. Although the need for federal approval or authorization limits to some degree the States' power to obtain information, requiring that approval has considerable advantages. The federal agencies are more likely to be informed about, and to understand, ERISA-related consequences and health care needs from a national perspective. Their involvement may consequently secure for the States necessary information without unnecessarily creating costly conflicts.⁶²

Federal agencies collecting data from self-insured plans and distributing it to the states is technically feasible, but it is a complex endeavor (discussed further below). However, many stakeholders and scholars believe that Justice Breyer's suggestion that the Department of Labor could "delegate" the authority to collect data to states is a promising alternative for agency action.⁶³

Breyer's opinion suggests two limits on the way such agency-based de-preemption must be structured, at least in his view. First, Breyer calls upon the agency to reflect an understanding of "ERISA-related consequences" of the action it is taking, which likely requires the Department of Labor to place at least some conditions on states' ability to collect data from group health plans. That is, blanket de-preemption of any state data collection efforts related to group health plans may not reflect an appropriately nuanced assessment of what is an appropriate requirement for group health plans under ERISA's preemption clause. Second, Breyer's concurrence envisions these data being available for use at the federal level. That is, if the information is being collected under the Department of Labor's

⁶¹ In a line of cases in which the Court has assessed whether the FDA's approval of and labeling requirements for a product preempts a state tort claim arising from injuries associated with that product, Breyer has emphasized the role that agency judgement can play in helping to understand the scope of preemption, and implicitly called upon agencies to provide that guidance. *Medtronic v. Lohr*, 518 U. S. 470 (1996) (Breyer, J., concurring in part and concurring in judgment), *Wyeth v. Levine* 555 U.S. 555 (2009) (Breyer, J., concurring). Cf. *Bates v. Dow Agrosciences LLC*, 544 U. S. 431 (2005) (Breyer, J., concurring) (addressing similar issues under the EPA's jurisdiction). But see [Sharkey].

⁶² *Gobeille v. Liberty Mutual Insurance Co.*, 577 U.S. ____ (2016) (Breyer, J., concurring).

⁶³ Sean Bland, Jeffrey Crowley, Lawrence Gostin, "Strategies for Health System Innovation After *Gobeille v Liberty Mutual Insurance Company*," 316 JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION 581 (August 9, 2016), <https://jamanetwork.com/journals/jama/article-abstract/2532230>; Nicholas Bagley, "A modest proposal for fixing *Gobeille*," *The Incidental Economist*, April 20, 2016, <https://theincidentaleconomist.com/wordpress/a-modest-proposal-for-fixing-gobeille/>; Maura Calsyn, "Policy Options to Encourage All-Payer Claims Databases," *Center For American Progress*, April 20, 2018, <https://www.americanprogress.org/issues/healthcare/reports/2018/04/20/449602/policy-options-encourage-payer-claims-databases/>; John Freedman, Linda Green, and Bruce Landon, "All-Payer Claims Databases — Uses and Expanded Prospects after *Gobeille*," *APCD Council*, December 2016, <https://www.apcdouncil.org/publication/all-payer-claims-databases-%E2%80%94-uses-and-expanded-prospects-after-gobeille/>; APCD Council, Comment Letter Regarding Employee Benefits Security Administration, Annual Reporting and Disclosure Proposed Rule, October 12, 2016, , <https://www.regulations.gov/document?D=EBSA-2016-0010-0046>.

authority to determine what is necessary or appropriate under ERISA and the ACA, then the federal government must have some ability to access it.

Breyer’s opinion thus suggests that the federal government could, for example, promulgate a regulation delegating to state APCDs the authority to collect claims data from group health plans, provided that the data is collected in the Common Data Layout using a routine electronic format for submission. As with legislative de-preemption, the Department of Labor could also consider limiting this delegation to data collection from plan administrators that are either responsible for at least a minimum number of self-insured lives in the state or have any insured business in the state. Similarly, the federal government could require APCDs that collect data to make the information available for research use, though that may require legislative changes in some states. Regardless of how other researchers access these data, federal agencies must reserve the ability to access the information themselves for enforcement or analysis purposes.

If the agencies pursued this approach, the authority underlying this regulation should follow the outline provided by the Court in *Gobeille*, invoking general ERISA authorities as well as the ACA. Specifically, as described above, ERISA requires plans to file an annual report,⁶⁴ and the Court majority explained that this language “permits the Secretary of Labor to ‘requir[e] any information or data from any [plan] where he finds such data or information is necessary to carry out the purposes of’” ERISA.⁶⁵ In addition, ACA section 1311(e)(3) describes a series of data elements related to plan enrollment and health care claims that plans offered through the Health Insurance Exchange must provide to their regulators, along with “other information as determined appropriate.”⁶⁶ Language codified into other federal statutes requires group health plans as well as other types of insured health benefits to provide the 1311(e)(3) data elements to their regulators, including the Department of Labor for group health plans.⁶⁷ But because the 1311(e)(3) data elements are required of many types of health plans – not just group health plans – the 1311(e)(3) authority may not, on its own, confer authority to delegate collection under ERISA to the states. The ERISA authority cited by the Court majority in *Gobeille* more clearly plays that role.

Create a National APCD

The preceding section considered policies that would restore something resembling the pre-*Gobeille* environment, with self-insured plans (and other payers) submitting data to independent state APCDs. These approaches could be implemented relatively quickly and would enable states to continue their ongoing work with complete data. In implementing this type of approach, Congress and federal agencies could take some steps to reduce administrative costs associated with submitting data to APCDs and perhaps modestly broaden access to these data. However, this approach will fall well short of addressing the other limitations of the existing patchwork of state APCDs, particularly the challenges in combining data from multiple states and the fact that many states lack APCDs. We turn now to approaches that would build a truly national APCD, with data collected from all payers and in all states.

A national APCD could, in principle, be implemented either legislatively or through agency action. As described above, the Supreme Court has suggested that existing statutes convey *authority* to the federal government to collect APCD-like data from insured and self-insured commercial health plans, and the federal government could directly furnish the national APCD with Medicare and Medicaid data. That said, collecting and maintaining these data would be a significant undertaking that would require an investment of funding and human capital. The required sums would likely not be large in the context of the federal health care budget, but building a national APCD would nonetheless be a major commitment requiring interagency effort from the Departments of Labor, Health and Human

⁶⁴ 29 U.S.C. 1024

⁶⁵ *Gobeille v. Liberty Mutual Insurance Co.*, 577 U.S. ____ (2016), quoting 29 U.S.C. § 1024(a)(2)(B).

⁶⁶ 42 U.S.C. § 13031(e)(3).

⁶⁷ 42 U.S.C. §§ 300gg–15a.

Services, and Treasury that would take resources away from other departmental priorities. Therefore, expansive federal data collection is most likely to occur if legislation mandates and funds such an effort. However, a motivated administration could likely undertake this initiative in the absence of Congressional action.

Below we describe the architecture of a national APCD, as we envision it. We describe the rules that would govern use of APCD data, privacy and security safeguards, funding requirements, as well as how the role of state APCDs would change following creation of a national APCD. We also address whether it would be preferable to house a national APCD in a non-governmental entity.

Architecture of a National APCD

The federal government would require all commercial payers to submit claims level data to a national APCD operated by a federal agency. Given the expertise of the Department of Health and Human Services (HHS) in maintaining Medicare claims and other health care data, it would be the best entity to actually possess and maintain this database. However, it would be appropriate for HHS to operate the project in consultation with the Department of Labor and the Department of Treasury, with which it shares jurisdiction over the requirements applied to health care payers.

Under this approach, all commercial payers – health insurance issuers, insured and self-insured group health plans, and non-federal governmental plans – that meet certain relatively low enrollment thresholds would be required to submit claims data. (Note that this differs from the scope of data collection contemplated by the Senate HELP committee, which required federal data collection only from self-insured payers.⁶⁸) Information would be submitted in a standard format specified by the agencies. Quarterly data submission may best balance policymakers’ and stakeholders’ desire for current and actionable information with minimizing the burden such data collection places on payers. Federal agency staff or contractors would be responsible for accepting data and performing the same types of quality checks that are conducted by state APCDs today and would have authority to enforce data submission requirements.

The federal government would furnish claims data for major public coverage programs to the APCD. The Centers for Medicare and Medicaid Services (CMS) directly holds claims data for traditional Medicare and collects similar data from Medicare Part D plans and Medicare Advantage plans.⁶⁹ CMS also collects claims-level data for state Medicaid programs and Children’s Health Insurance Programs via its Transformed Medicaid Statistical Information System (T-MSIS), although there are currently some questions about the quality of the T-MSIS data.⁷⁰ If those concerns persist, federal policymakers could consider instead obtaining Medicaid data via agreements with state APCDs, as we discuss below.

The Office of Personnel Management also likely has authority to facilitate collection of data from the Federal Employees Health Benefit Program. Since 2011, the agency has been investing in efforts to establish its own claims database; some carriers have opposed these efforts, but the agency insists it has legal authority to collect these data for its own purposes and for independent research.⁷¹ Using the

⁶⁸ See Lower Health Care Costs Act, S. 1895, 116th Congress § 303 (2019), <https://www.congress.gov/bill/116th-congress/senate-bill/1895>.

⁶⁹ The encounter data collected from Medicare Advantage plans is currently believed to have some data quality problems. See “Report to the Congress: Medicare and the Health Care Delivery System,” *MedPac*, June 2019, http://www.medpac.gov/docs/default-source/reports/jun19_ch7_medpac_reporttocongress_sec.pdf?sfvrsn=0. Ensuring that the national APCD had complete and accurate data would be another reason to continue efforts to improve the quality of these data.

⁷⁰ “Update on Transformed Medicaid Statistical Information System (T-MSIS),” *Medicaid and CHIP Payment and Access Commission*, October 2019, <https://www.macpac.gov/publication/update-on-transformed-medicaid-statistical-information-system-t-msis/>.

⁷¹ See, “Privacy Impact Assessment for the Health Claims Data Warehouse (HCDW),” *Office of Personnel Management Office of Healthcare and Insurance*, May 8, 2018, <https://www.opm.gov/information-management/privacy-policy/privacy-policy/hcdw.pdf>; “Congressional Budget Justification and Annual Performance Plan,” *U.S. Office of Personnel Management*,

same authority, policymakers would fold the current effort into the national APCD, which would likely reduce burdens on insurers since carriers submitting to the OPM database would likely be submitting data to the APCD for their other products. The Veterans Health Administration could likely provide its data to the national APCD as well. The Department of Defense maintains similar data for its workforce in the TRICARE Encounter Data system and provides some limited access to civilian researchers.⁷² While similar to the information contained in an APCD, these data could have national security implications that limit the degree to which they can be shared; these issues are beyond the scope of this paper, but we believe the Department of Defense could likely share some limited data with a national APCD.

For all payers, we anticipate that the APCD would collect information similar to the information held by existing state APCDs, including the standard fields included on health care claims, patient demographic information, and certain plan characteristics.

Uses of APCD Data

We envision that the data held by a national APCD would be used for public reporting, research, and policymaking, like data held by state APCDs. In particular, agencies could produce routine reports on health care utilization, prices, and spending, as well as dimensions of health care quality that can be measured in claims data, both nationwide and disaggregated by geography. These reports would offer data users a sharper and more detailed picture of national trends, and the large sample sizes would particularly improve the ability to compare geographic areas. Agencies would also be able to conduct narrower analyses linked to current national priorities, just as state APCDs have produced analyses related to topics of current public and policy interest, such as the opioid epidemic.

Data in a national APCD would also be used to directly support policy design and implementation. Legislative agencies like the Congressional Budget Office, Government Accountability Office, Medicare Payment Advisory Commission, and Medicaid and CHIP Payment and Access Commission would access these data to conduct policy analysis and oversight through mechanisms similar to those they use to access Medicare and Medicaid data today. Executive branch policymakers could also use the data to conduct analyses to inform policy deliberations and, where relevant, for policy implementation. Use of APCD data for law enforcement or immigration enforcement purposes would be prohibited to avoid any risk of discouraging individuals from seeking appropriate health care (particularly care for substance use disorders), with narrow exceptions for anti-trust enforcement and investigations of health care fraud.

The national APCD would also make its data available to researchers. We expect that researchers could access APCD data in a manner similar to the way they access Medicare and Medicaid claims data today. Basic public use files that strip out all potentially identifying information could be made available through a simple process, and researchers could apply for access to more detailed data sets, subject to stringent privacy protections as described below. The existing Research Data Access Center (ResDAC)⁷³ that helps researchers apply for and use Medicare and Medicaid data could be expanded to support researchers seeking data from the national APCD.

February 2018, <https://www.opm.gov/about-us/budget-performance/budgets/congressional-budget-justification-fy2019.pdf>; Arthur Allen, "Insurers' Doubts Idle OPM Data Warehouse," *Politico*, December 13, 2017, <https://www.politico.com/story/2017/12/13/insurers-doubts-idle-opm-data-warehouse-294976>.

⁷² Office of the Assistant Secretary of Defense for Health Affairs (OSAD(HA)), TRICARE Management Activity (TMA), Human Research Protection Program (HRPP), "Guide for DoD Researchers on Using MHS Data," October 10, 2012, <https://health.mil/Reference-Center/Publications/2012/10/10/Guide-for-DoD-Researchers-on-Using-MHS-Data>.

⁷³ Research Data Assistance Center, "Find, Request and Use CMS Data," <https://www.resdac.org/> (last visited September 18, 2020).

A national APCD would need to develop rules about data access for organizations that may sell data products derived from APCD data. Medicare currently allows certain “Qualified Entities” (QE) to obtain Medicare data and sell products based on that data, provided that they also combine Medicare data with commercial claims data to produce certain public reports.⁷⁴ We believe that there is no reason to bar commercial entities from accessing data and packaging it in ways that may be valuable for downstream users and that a national APCD would ideally create a counterpart to the QE program that enables access to national APCD data, subject to appropriate requirements.

Across all of these uses, policymakers would need to decide whether users of APCD data would be permitted to publicly disclose provider- or payer-level estimates, particularly estimates of negotiated prices. Many state APCDs permit public disclosure of provider-or payer-level data, and the Trump Administration has recently proposed several policies intended to make health care prices more transparent, including requiring providers to make their negotiated prices public.⁷⁵

The main potential advantage of allowing these types of disclosures is that it could support research on natural experiments that involve specific providers and insurers, which can provide insights about health care market dynamics that would otherwise be unavailable.⁷⁶ However, provider and payer stakeholders are likely to oppose such disclosures through an APCD. Further, as discussed earlier, some believe that disclosure of negotiated prices could put upward pressure on prices, which would be both substantively undesirable and, as a procedural matter, could lead the Congressional Budget Office to estimate that legislation creating a national APCD that allowed such disclosures would increase federal spending. That said, as also discussed above, other evidence suggests that greater price transparency may not meaningfully increase prices or may even put modest downward pressure on prices. On balance, we lean toward permitting APCD users to report provider- and payer-specific estimates but acknowledge that there are arguments against a permissive approach.

Role of Existing State APCDs

If the federal government created a national APCD, the role of state APCDs might change. In particular, we envision the federal government would share with a state APCD all data collected from that state shortly after federal receipt, including data for self-insured plans that state APCDs cannot collect today.⁷⁷ While states could continue collecting data from payers themselves, we anticipate that few state APCDs would choose to do so, provided that the federal government adopted appropriate quality control processes and prioritized delivering data to states in a timely fashion.⁷⁸ To make ceasing data collection more attractive for states, the federal government could allow state APCDs to direct insured payers to provide some limited state-specific data elements to the national APCD, which would then be part of the data the federal government provided back to the state.

⁷⁴ “Qualified Entity Program,” *U.S. Centers for Medicare & Medicaid Services*, October 15, 2019,

<https://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/OEMedicareData>.

⁷⁵ See, e.g., U.S. Department of Health & Human Services, “Trump Administration Announces Historic Price Transparency Requirements to Increase Competition and Lower Healthcare Costs for All Americans,” November 5, 2019,

<https://www.hhs.gov/about/news/2019/11/15/trump-administration-announces-historic-price-transparency-and-lower-healthcare-costs-for-all-americans.html>.

⁷⁶ See, e.g., Glenn Melnick and Katya Fonkych, “An Empirical Analysis of Hospital ED Pricing Power,” 26 *AMERICAN JOURNAL OF MANAGED CARE* 105 (December 19, 2019), <https://www.ajmc.com/journals/issue/2020/2020-vol26-n3/an-empirical-analysis-of-hospital-ed-pricing-power>; Mark Shepard, “Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange,” *Harvard Kennedy School of Government*, August 1, 2016, https://scholar.harvard.edu/files/mshpard/files/mshpard_hospitalnetworksselection_Aug2016.pdf.

⁷⁷ This approach, as well as the possible approach to Medicaid collection described in the next paragraph, parallels the approach envisioned for self-insured data under the Senate HELP Committee’s Lower Health Care Costs Act.

⁷⁸ Medicaid data is a possible exception, as noted above. If T-MSIS data were determined to be inadequate, state APCDs could continue to obtain Medicaid data from their state Medicaid agencies and provide those data to the national APCD in exchange for the data collected by the national APCD (paralleling the structure envisioned in the recent Senate HELP bill).

Over the long run, centralizing data collection would reduce states' data collection costs, allowing them to focus their resources on data analysis and policy support. In the near term, however, the creation of a national APCD would impose some transition costs. Receiving data from the national APCD (rather than receiving it directly from each payer) would require state APCDs to develop processes for accepting and integrating that information into their data systems. It may also require changes in state law to align privacy or other standards.⁷⁹ Note that, in contrast to the HELP Committee's bill, we envision a process where the national APCD would collect all data, not just data for self-insured plans. In addition to better facilitating use of the data by federal policymakers and multi-state analyses, we believe this approach ultimately reduces burden for state APCDs because the HELP process would require states to build the ability to accept national data without relieving them of the need to maintain their own data collection systems.

States would lose some control in the shift to national data collection. While they would retain the ability to collect limited state specific elements and could likely adapt to the standardized data format, they would not be able to direct data submission from small entities or mandate certain formatting. Nor would they be able to oversee submission, conduct their own quality control processes, or leverage their in-state relationships to promote timely and accurate compliance with reporting standards. That said, we expect a national APCD could achieve the same – or better – levels of overall data quality through a national quality control process and clearly articulated federal penalties for noncompliance.

Privacy and Security Safeguards

A national APCD would be powerful because it contains detailed information about health care delivery, including who received which health care services, who delivered those services, and who paid for them. But those data are obviously sensitive and, as noted earlier, some observers have expressed concern that holding claims-level information in a federal database poses risks to privacy. An important question, therefore, is how to ensure the privacy and security of data held in a national APCD.

To start, we note that while a national APCD is a new undertaking, it would not present fundamentally novel privacy or security concerns. As noted above, the federal government already possesses large amounts of claims data through operation of the Medicare and Medicaid programs, and, while there have been isolated security incidents, we are unaware of any significant data breach resulting from the use of claims data for program operations, public reporting, or research, a notable contrast with some private payers.⁸⁰ Consistent with this, the privacy and security of data in a national APCD can be ensured by adapting the procedures that the federal government already uses to safeguard claims data.

The starting point should be to ensure that the APCD is subject to the Privacy Act, which protects the privacy and security of personal information held by the government, as well as the privacy and security requirements that apply to covered entities under the Health Insurance Portability and Accountability Act (HIPAA), which protect health information generally. The former would likely be automatic, while making an APCD subject to HIPAA rules could require explicit action by policymakers.⁸¹

Tying an APCD's privacy and security safeguards to these existing federal laws would be superior to creating a new legal framework to govern privacy and security for the APCD. These existing frameworks have generally been successful in protecting health information held by the federal government. Indeed, the Privacy Act and HIPAA are the principal laws that govern CMS' handling of

⁷⁹ See APCD Council, on behalf of State APCDs, Letter to the U.S. Senate Committee on Health, Education, Labor and Pensions regarding the Lower Health Care Costs Act of 2019, June 5, 2019, https://www.apcdouncil.org/sites/default/files/media/lowerhealthcarecostsact_comments_06052019_final.pdf.

⁸⁰ See, e.g., Nate Lord, "Top 10 Biggest Healthcare Data Breaches of All Time," *Data Insider*, June 25, 2018, <https://digitalguardian.com/blog/top-10-biggest-healthcare-data-breaches-all-time>.

⁸¹ In particular, it is unclear whether an APCD would be a "covered entity" within the meaning of HIPAA and, thus, whether HIPAA's privacy and security rules would automatically apply to an APCD.

Medicaid and Medicare claims data, and, as noted above, we are unaware of any significant data breaches affecting these claims databases. Linking protections for APCD data to these existing statutes also takes advantage of the fact that these statutes, as well as their accompanying guidance and regulations, are periodically updated in response to changing technology and other developments.

We envision that—pursuant to these laws—an APCD would adopt procedures to prevent inappropriate disclosure similar to those CMS uses to control access to Medicare and Medicaid claims data. For example, all researchers seeking identifiable data from CMS (that is, data from which an individual’s identity could potentially be discerned) must sign a data use agreement in which they agree to abide by specified security requirements and agree not to release results pertaining to groups of people smaller than 11.⁸² Researchers seeking datasets that contain the largest array of identifiable data elements must additionally submit a detailed application describing why their research project requires identifiable data, and the study must pass human subjects review by an Institutional Review Board operating under the Common Rule as well as review by the CMS Privacy Board.⁸³ Many studies using CMS data now access and analyze those data via CMS’ Virtual Research Data Center, a secure computing environment maintained by CMS, rather than by receiving the data files directly, which allows CMS to retain control over the data even when used by researchers; a similar approach could be used in the context of a national APCD. Similar processes apply to other non-CMS users, including state governments and other federal agencies.⁸⁴

We believe these procedural safeguards aimed at preventing inappropriate disclosure are the most important part of efforts to protect the privacy and security of information held by an APCD. However, as a further step to address privacy and security concerns, policymakers could consider limiting the APCD’s ability to collect or retain identifiable data. In considering options like these, it is useful to distinguish between two types of identifiable data elements:

- **Direct identifiers.** Health care claims data contain some fields that directly identify patients, such as a patient’s name or social security number.⁸⁵ Completely barring the APCD from interacting with direct identifiers would make it impossible to use the APCD for longitudinal analyses that follows patients over time as they are served by different providers and covered by different insurers, which would substantially limit the questions an APCD could answer. For example, being unable to follow patients would make it impossible to use an APCD to study the care patients receive over the course of a pregnancy or in connection with a chronic disease.

However, it is possible to facilitate longitudinal analyses without retaining direct identifiers within the APCD. In particular, it is possible to use direct identifiers in the original data to create an “encrypted” unique identifier that links together different records corresponding to the same person but does not itself reveal any identifying information. Indeed, with rare exceptions, CMS only provides encrypted unique identifiers when making Medicare and Medicaid data available for analytic use, and many other entities that hold identifiable data follow similar practices.⁸⁶ An APCD would almost surely follow similar practices when making

⁸² “Limited Data Set (LDS) Files,” *U.S. Centers for Medicare & Medicaid Services*, June 30, 2020, https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS.

⁸³ Research Data Assistance Center, “Research Identifiable File (RIF) Requests,” <https://www.resdac.org/research-identifiable-files-rif-requests> (last visited September 18, 2020).

⁸⁴ Research Data Assistance Center, “Research Identifiable File (RIF) Requests,” <https://www.resdac.org/research-identifiable-files-rif-requests> (last visited September 18, 2020); “Identifiable Data Files,” *U.S. Centers for Medicare & Medicaid Services*, March 20, 2020, <https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/Researchers>.

⁸⁵ HIPAA regulations define a longer list of fields considered to be direct identifiers. See 45 CFR § 164.514(e).

⁸⁶ See United States Census Bureau, “Data Ingest and Linkage,” September 7, 2016, <https://www.census.gov/about/adrm/linkage/technical-documentation/processing-de-identification.html> and the cited technical paper for a discussion of how the Census Bureau creates this type of encrypted identifier to enable research using the various identifiable databases it holds.

data available for analytic purposes, but policymakers could go one step farther and bar the APCD from even *holding* anything other than an encrypted identifier on an ongoing basis.

This type of approach could allay some stakeholder concerns and would have the advantage of making the APCD a less attractive target for identity theft or other large-scale fraudulent activities. However, even this restriction would have some downsides. In particular, direct identifiers in real-world claims data often contain minor errors and imperfections, like transposed digits, which can complicate creation of a unique identifier. Completely forbidding an APCD from holding unencrypted direct identifiers would limit the APCD's ability to investigate approaches to creating encrypted identifiers that are more robust to data errors or to take advantage of improvements in methods for creating encrypted identifiers.

- **Indirect identifiers.** Many fields on health care claims databases that do not directly identify patients, such as dates of service, patient age, and patient zip codes, can nevertheless be used to identify individual patients when used in combination with each other and with other fields present on a health care claim.⁸⁷ For example, claims data reflecting trauma care in a specific city on a specific date could be linked to news reports. However, preventing an APCD from collecting and retaining these types of indirect identifiers would severely limit the types of research an APCD could support. To take just one timely example, barring the APCD from holding indirect identifiers would likely prevent an APCD from simultaneously holding fine-grained data on both patient age and patient zip code. That, in turn, would reduce the usefulness of an APCD for studying COVID-19 due to the large differences in the pandemic's impact by age and geography.

The claims database that would have been created by the recent Senate HELP proposal would have permitted the non-profit operating the database to collect identifiable information but would have required the non-profit to subsequently de-identify those records.⁸⁸ Consistent with the discussion above, this process would allow creation of an encrypted unique identifier and, thus, facilitate use of the database for longitudinal analysis. However, it could require the removal of many data elements that may be indirect identifiers (like zip code or age), seriously limiting the database's capabilities.

Funding Requirements

We have not produced a detailed estimate of what a national APCD might cost, but similar federal undertakings can provide some guidance on this question. Notably, the Agency for Healthcare Research and Quality (AHRQ) operates the Healthcare Cost and Utilization Project (HCUP), which collects data from state-operated inpatient and outpatient hospital encounter databases and then makes harmonized versions of those databases available to researchers for a fee.⁸⁹ It also produces periodic reports based on its data and provides a web-accessible tool that can be used to produce aggregate tabulations without purchasing the underlying discharge databases. Budget documents show that funding for HCUP was between \$9 and \$14 million in fiscal year 2020.⁹⁰

⁸⁷ See, e.g., Gregory Simon, Susan Shortreed, Yates Coley, Robert Penfold, Rebecca Rossom, Beth Waitzfelder, Katherine Sanchez, and Frances Lynch, "Assessing and Minimizing Re-identification Risk in Research Data Derived from Health Care Records," 7 THE JOURNAL FOR ELECTRONIC HEALTH DATA AND METHODS 6 (March 29, 2019), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6450246/>; Khaled El Emam, Elizabeth Jonker, Luk Arbuckle, and Bradley Malin, "A Systematic Review of Re-Identification Attacks on Health Data," 10 PLOS ONE 4 (December 2, 2011), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3229505/>.

⁸⁸ Lower Health Care Costs Act, S. 1895, 116th Congress § 303 (2019), <https://www.congress.gov/bill/116th-congress/senate-bill/1895>.

⁸⁹ "HCUP Fact Sheet," *Healthcare Cost and Utilization Project*, https://www.hcup-us.ahrq.gov/news/exhibit_booth/HCUPFactSheet.pdf (last visited September 17, 2020).

⁹⁰ See "National Institute for Research on Safety and Quality (NIRSQ)," *Department Of Health And Human Services National Institutes Of Health*, https://www.ahrq.gov/sites/default/files/wysiwyg/cpi/about/mission/budget/2021/FY_2021_CJ_NIRSQ.pdf (last visited September 18, 2020). The budget request indicates that AHRQ funding for HCUP and several other activities totaled \$14.3 million in fiscal year 2020, which places an upper bound on the amount spent on HCUP. The request also indicates that the

The analogy between a national APCD and HCUP is imperfect. A national APCD would receive data from a greater number of entities than HCUP and would receive and manage a greater variety of types of data. A national APCD would also need to do all of its own data quality checks, whereas HCUP is able to rely to some degree on data quality checks performed by its state partners. On the other hand, a national APCD would collect data through a uniform process and format, whereas each of HCUP's state partners submits data in slightly different form, which could reduce the effort a national APCD would need to invest in harmonizing different data sources.

On balance, we expect that operating a national APCD would be a more complex undertaking than operating HCUP, though not overwhelmingly so. Correspondingly, a reasonable guess might be that a national APCD would cost around twice what it costs to operate HCUP or around \$20 million per year. Costs are likely to be higher initially as the APCD creates its basic data systems, so policymakers would likely need to provide additional start-up funding, perhaps on the order of \$40 million. These estimates are consistent with the evidence described above that reveal state APCD operating budgets (operating on a smaller scale with fewer payers) of \$1.6 to \$4.4 million. Similarly, the Office of Personnel Management spent approximately \$10 million to develop a claims database for the FEHB program.⁹¹ Further analysis to refine these cost estimates would be worthwhile.⁹²

An Alternative Governance Structure: Housing a National APCD Within a Non-Profit

Some stakeholders have expressed concern about creating a federal database that holds detailed claims information. While some of those concerns reflect questions about how policymakers would maintain the security of APCD data, which we discussed above, some stakeholders may also harbor a general uneasiness about the government, particularly the federal government, possessing such a large quantity of health information, separate from fears regarding a potential breach.

As noted above, it is not fundamentally novel for a government entity to hold these types of data; the federal government holds claims data for the Medicare and Medicaid programs, and state APCDs will continue to exist and possess similar data in the absence of federal legislation. Nevertheless, given stakeholder concerns, policymakers have considered an alternative where Congress would direct that the national APCD data be collected and maintained by a non-profit organization under contract with the federal government.⁹³

Under this type of approach, a non-profit entity (likely an existing organization that has experience with health care claims data) would receive a time-limited contract with the federal government. As a contractor, they would receive claims data from payers, make those data available to a variety of authorized users, and conduct their own research. A board of experts and stakeholders would provide oversight and advice on the maintenance and handling of the data. While contractors would almost certainly be involved in supporting any APCD that was housed within a federal agency, this type of approach differs by providing more autonomy and direct responsibility to the contractor.

Administration's fiscal year 2021 proposal to provide \$8.8 million for HCUP alone would only partially fund HCUP, which provides a lower bound on the amount spent on HCUP.

⁹¹ Arthur Allen, "Insurers' Doubts Idle OPM Data Warehouse," *Politico*, December 13, 2017, <https://www.politico.com/story/2017/12/13/insurers-doubts-idle-opm-data-warehouse-294976>.

⁹² CMS also collects claims data and other similar data from health insurers to operate risk adjustment programs and for other similar purposes, including from insurers in the Medicare Part D program, the Medicare Advantage program, and the individual and small group markets. These data collection functions performed by these systems are, in most important respects, closely analogous to the functions that would need to be performed by a national APCD. Unfortunately, public estimates of spending on those systems is not readily available because budget documents combine spending on these data systems with other agency activities. However, the cost of those data systems would provide a useful datapoint for estimating the cost of a national APCD.

⁹³ See Lower Health Care Costs Act, S. 1895, 116th Congress § 303 (2019), <https://www.congress.gov/bill/116th-congress/senate-bill/1895>.

The framework poses important governance challenges, though steps can be taken to mitigate these difficulties to some degree. Specifically, three strategies can help ensure that the entity remains accountable to federal policymakers and the public interest:

- **Flexibility to change contractors.** Authorizing legislation and agency contracts should ensure that the federal government retains the ability – in practice, and not just in theory – to change contractors in the face of poor performance. The federal government should retain ownership of software products and require transition assistance in its contracts, and legislation should be drafted broadly to ensure multiple entities could be viable contractors.
- **Policymaker access to data.** Authorizing legislation and agency contracts should ensure that executive and legislative agency staff have flexible access to the data and are able to pursue agency objectives without interference from the contractor.
- **Clear federal control.** While a stakeholder board can provide some additional oversight of contractor performance, it is important that the federal government itself retain the authority to supervise the contractor and hold it accountable in the event of poor performance. Stakeholders should not gain the ability to direct research away from areas that affect their commercial interests.

In addition to potentially offering a more politically appealing path, a contractor-led approach may offer additional flexibility and agility in research. Contractors operating with significant autonomy could be effective in quickly developing usable data products that reflect emerging interests. On the other hand, even with governance safeguards, the inherent difficulties in holding an outside entity accountable raises the risk of poor performance.

Harmonize State APCDs and Create a Federal Clearinghouse for APCD data

If building a national APCD is judged infeasible or undesirable, a less ambitious approach would be to attempt to “stitch together” the current patchwork of state APCDs in ways that can overcome some—though not all—of the shortcomings of relying on a network of state APCDs. In many respects, this approach would mirror the approach the federal government has taken to stitch together state hospital encounter databases via the Healthcare Cost and Utilization Project (HCUP) operated by the Agency for Healthcare Research and Quality (AHRQ). Like creating a national APCD, it would likely be feasible to pursue this approach through agency action, but this type of project would be most likely to succeed if Congress mandates and funds the effort.

This approach would involve four main steps:

- **Facilitate state collection of data from self-insured plans.** To ensure that the state APCDs are able to provide a comprehensive picture of their commercial insurance markets, the federal government would grant state APCDs the authority to collect data from self-insured plans if they provided data to the federal clearinghouse described below. This could occur via either the legislative or administrative pathways described earlier.
- **Provide performance-contingent grants to state APCDs.** The federal government would provide grant funding to state governments to support the creation and maintenance of APCDs, with two objectives. First, the grant funding would encourage states that do not currently operate APCDs to set them up.⁹⁴ Second, the funding would allow the federal

⁹⁴ The federal government has had some success in using grant funding to encourage states to invest in data collection efforts. See, e.g., U.S. Centers for Medicare & Medicaid Services, “Rate Review Cycle III Funding Opportunity: Frequently Asked Questions,” September 18, 2020, <https://www.cms.gov/CCIIO/Resources/Fact-Sheets-and-FAQs/rr-foa-faq-6-6-2013>.

government to place certain requirements on how state APCDs collect and share data, as described below.

While it might be possible to use existing funding to support this type of grant program, ideally this grant program would be created and funded legislatively. To allow states to make long-term plans and investments, legislation should ideally provide a permanent mandatory appropriation. In light of the data on typical state APCD budgets discussed earlier, a reasonable estimate is that grant funding on the order of \$2 million per state per year would be adequate to encourage state APCDs to comply with federal requirements. Inducing new states to set up APCDs might require larger amounts, as discussed below.

- **Create a federal clearinghouse for APCD data.** State APCDs that accept federal grant funds or wish to collect data from self-insured plans would be required to report the data they collect (including data reported by fully-insured plans) to the federal government, which would then integrate the various states' data with federal Medicare and Medicaid data in a single harmonized database.⁹⁵ As noted earlier, HCUP successfully performs a similar function with respect to state hospital encounter databases. Indeed, policymakers could consider making the federal clearinghouse part of the broader suite of HCUP databases.

As with a national APCD, we anticipate that the federal government would use the harmonized database to produce public reports and make the database available to researchers and policymakers. Similarly, we anticipate that the database would abide by privacy and security safeguards similar to those we envision for a national APCD.

The federal government would need to commit meaningful resources to support this type of clearinghouse. While we have not developed a formal cost estimate, the \$9-14 million per year that the federal government currently spends on HCUP provides a reasonable point of comparison. Relative to HCUP, the clearinghouse would need to manage a greater variety of types of data but would have somewhat greater control over how that data is collected and submitted. On balance, we suspect that operating the clearinghouse would be modestly more complex than operating HCUP. Correspondingly, while a reasonable guess is that operating the clearinghouse might cost around \$15 million per year on an ongoing basis, spending needs would likely be higher initially, and further analysis would be worthwhile. Some resources might be available in existing funding streams, but ideally Congress would pass new legislation directing the federal government to pursue this project and appropriating the needed funds.

- **Set common data collection standards.** To maximize the utility of the federal clearinghouse, the federal government would need to require state APCDs to abide by certain minimum data collection standards (for both insured and self-insured plans). Those standards would need to specify the minimum set of data elements states are required to collect, the schedule on which states would be required to collect and submit data, and a set of data quality standards that states would be expected to meet. Indeed, one limitation on the HCUP databases has been that some states' discharge databases do not collect certain data elements or do not collect those data elements in comparable ways, which can complicate multi-state research projects.⁹⁶ Federal policymakers could simultaneously seek to standardize how state APCDs collect data from payers, including by setting standards for the submission format and process akin to those discussed in the context of legislative and agency de-preemption. Compliance with all of these data collection standards could be made a condition of the grant

⁹⁵ As under a national APCD, if data quality concerns with the T-MSIS data held by CMS persisted, the clearinghouse could consider instead collecting those data via the state APCD.

⁹⁶ See Healthcare Cost and Utilization Project, "Availability of Data Elements by Year," August 2, 2019, https://www.hcup-us.ahrq.gov/db/state/siddist/siddist_ddeavailbyyear.jsp.

funds described above or of granting states the authority to collect data from self-insured plans.

This approach would be a substantial improvement over the status quo. Notably, it would restore the comprehensiveness of state APCDs, make it much easier to combine data from multiple states to support public reporting, research, and policymaking, as well as ensure that federal policymakers have ready access to APCD data.

However, relative to creating a national APCD, this approach would have some important limitations. First, while the grant funding we envision under this approach might encourage some additional states to create APCDs, it is unlikely to motivate all states to overcome the political opposition that APCDs can engender. Second, even with the common data collection standards envisioned above, it is unlikely that data would be perfectly comparable across states with so many different entities responsible for data collection. Third, harmonizing data submission process in the ways envisioned above would likely only modestly reduce administrative burdens for payers required to submit to multiple states, and it would do essentially nothing to reduce duplication of state APCD infrastructure.

Policymakers could, in principle, address the first of these problems (incomplete coverage) by allowing the federal government to operate a federal APCD in states that decline to set up an APCD or that wish to cede these functions to the federal government, essentially creating a hybrid of the clearinghouse approach envisioned in this section and the national APCD approach discussed earlier. This approach could ensure truly national coverage and create a platform that could encourage migration toward a truly national APCD over the long run, albeit at higher cost at least in the short run.

The recent Senate HELP Committee proposal (see Box 2) offered a different form of hybrid approach, with some advantages and disadvantages relative to the hybrid approach discussed in the last paragraph. On the positive side of the ledger, the HELP bill envisioned the federal government handling all data collection related to self-insured plans, which would likely both improve data quality and reduce administrative burden. However, the HELP bill had no mechanism to collect insured data in states without APCDs, and it is unclear whether the HELP bill would have provided the authority required to regulate state APCDs' data collection practices as we envision above.

A Note on “Federated” Alternatives to APCDs

As an alternative to the policy approaches considered in this section, particularly creating a national APCD, some health plans have suggested creating a “federated” or “distributed” claims data system.⁹⁷ Under this approach, each plan would retain possession of its own data, but data users could query those plan-specific databases under certain circumstances. We are unaware of any fully fleshed-out proposal to create a federated system as an alternative to an APCD. However, advocates of this approach seem to have two broad architectures in mind, each of which we discuss in turn.

Under the first architecture, payers would transmit only aggregate summary statistics in response to queries from data users.⁹⁸ This structure mirrors the External Data Gathering Environment (EDGE)

⁹⁷ See, e.g., “Achieving States’ Goals for All-Payer Claims Databases,” *Anthem Public Policy Institute*, June 2018, https://www.antheminc.com/cs/groups/wellpoint/documents/wlp_assets/d19n/mzq1/~edisp/pw_g345393.pdf; Sheryl Turney, “Claims-based Databases for Policy Development and Evaluation: Testimony before the National Committee on Vital and Health Statistics,” June 17, 2016, <https://www.ncvhs.hhs.gov/wp-content/uploads/2016/05/Panel-3-Sheryl-Turney-Anthem-2016June17.pdf>; Joel Slackman, “Hearing On Claims-Based Databases For Policy Development and Evaluation: Testimony before The National Committee On Vital And Health Statistics,” June 17, 2016, <https://ncvhs.hhs.gov/wp-content/uploads/2016/05/Panel-1a-Joel-Slackman-BCBSA-Written-2016June17-without-password.pdf>; Blue Cross Blue Shield Association. “Ensuring the Privacy and Security of Patient Data Distributed Secure Access Data Model,” (on file with authors).

⁹⁸ Joel Slackman, “Hearing On Claims-Based Databases For Policy Development and Evaluation: Testimony before The National Committee On Vital And Health Statistics,” June 17, 2016, <https://ncvhs.hhs.gov/wp-content/uploads/2016/05/Panel-1a-Joel-Slackman-BCBSA-Written-2016June17-without-password.pdf>.

servers used to collect data from payers for the individual and small group market risk adjustment programs.

This type of system would be dramatically less useful to data users than a traditional APCD. Most importantly, because end users would receive only summary statistics from each contributing payer's data, end users would need to write specialized code to compute market-wide aggregates, rather than calculating such amounts using standard statistical packages. For anything more complicated than calculating market-wide means (e.g., fitting a regression model), this would be a formidable task. And some analyses would be effectively impossible under this structure because many statistics—including statistics as simple as the market-wide median payment for a particular service—cannot be calculated based solely on summary statistics from each payer's data. It is also often hard to assess data quality and modify analyses to mitigate data quality problems without access to claims-level information, which could threaten the reliability of analyses performed using this type of system.

Advocates of this approach argue that it would better protect the security of claims data. Indeed, data users would no longer be able to access to claims-level information, which would remove one potential source of a breach; however, as discussed earlier, those risk can be mitigated in other ways. And beyond limiting data users' access, the security advantages of this approach are unclear. Each payer would now need to set up an internet-connected server that contains its claims data and responds to queries from data users, rather than submitting data once to the APCD and having the APCD handle interactions with end users. Thus, the number of potential sources of a data breach would be much higher under this type of arrangement, although the number of records exposed in any given breach would be smaller.

Under the second architecture, data users could obtain claims-level information from each payer and assemble a temporary local dataset for analysis.⁹⁹ This approach would avoid many of usability pitfalls of the EDGE-like approach described above, but it would still have important weaknesses relative to a traditional APCD; notably, data users could not benefit from the APCD's data curatorial efforts, particularly efforts to identify and resolve data quality problems and ensure comparability of data elements across data submitters. Moreover, this structure would have no meaningful security advantages relative to a traditional APCD and would arguably be worse in light of the fact that, as under the EDGE-like model, each payer would need to host its claims data on its own internet-connected server.

A final important note is that, regardless of the precise architecture, many supporters of federated approaches appear to envision that each individual payer would approve or disapprove use of its data on a case-by-case basis.¹⁰⁰ Indeed, this may be the key feature of a federated approach from the perspective of its proponents. However, given the large number of payers involved, requiring payer approval would likely make using these data prohibitively burdensome in most applications. It would also inappropriately constrain research that payers viewed as opposed to their parochial interests.

Conclusion

APCDs are important tools for understanding and improving our health care system, but existing APCDs have major limitations. State APCDs' inability to collect data from self-insured plans prevents them from providing a complete picture of health care enrollment and payment within a state. Moreover, the 23 states that have APCDs today encompass only half the population, and our existing

⁹⁹ Blue Cross Blue Shield Association. "Ensuring the Privacy and Security of Patient Data Distributed Secure Access Data Model," (on file with authors).

¹⁰⁰ "Achieving States' Goals for All-Payer Claims Databases," *Anthem Public Policy Institute*, June 2018, https://www.antheminc.com/cs/groups/wellpoint/documents/wlp_assets/d19n/mzq1/~edisp/pw_g345393.pdf; <https://ncvhs.hhs.gov/wp-content/uploads/2016/05/Panel-1a-Joel-Slackman-BCBSA-Written-2016June17-without-password.pdf>; Joel Slackman, "Hearing On Claims-Based Databases For Policy Development and Evaluation: Testimony before The National Committee On Vital And Health Statistics," June 17, 2016, <https://ncvhs.hhs.gov/wp-content/uploads/2016/05/Panel-1a-Joel-Slackman-BCBSA-Written-2016June17-without-password.pdf>.

patchwork system makes it difficult to conduct analyses with multiple states' data, limits the availability of these data for federal policymaking, and fails to exploit potential economies of scale.

Federal policymakers have a variety of options to redress these problems. They could: (1) act surgically to undo the effects of the 2016 Supreme Court decision that blocks states from collecting APCD data; (2) build a truly national APCD covering all states and all payers; or (3) work to harmonize existing state APCDs and encourage states that currently lack APCDs to create them.

We believe that creating a national APCD is the best of these approaches. With a small investment of resources relative to national health care spending, policymakers could create a tool that offers a comprehensive picture of the health care system that no existing data source can come close to providing, thereby accelerating efforts by a wide range of public and private actors to better understand and ultimately improve American health care. Efforts to expand state APCD coverage and harmonize existing APCDs could also improve on the status quo (as well as the pre-*Gobeille* status quo) and might encounter somewhat less stakeholder opposition. That said, such a project would achieve less than creating a truly national APCD, and we view it as a decidedly second-best alternative.

Finally, we underscore that if the political will cannot be summoned to pursue these larger projects, either Congress or federal agencies should act swiftly to at least restore states' ability to collect self-insured data. While not costless, this is a fairly simple undertaking that will make existing APCDs far more useful and comprehensive. It would also allow state APCDs to maximize their own potential, hopefully building a constituency for creating a better national infrastructure over the longer-term.



The USC-Brookings Schaeffer Initiative for Health Policy is a partnership between the Economic Studies Program at Brookings and the USC Schaeffer Center for Health Policy & Economics and aims to inform the national health care debate with rigorous, evidence-based analysis leading to practical recommendations using the collaborative strengths of USC and Brookings.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.