BlueCross BlueShield
Association

An Association of Independent
Blue Cross and Blue Shield Plans

1310 G Street, N.W.
Washington, DC 20005
www.BCBS.com

TESTIMONY

Before the

NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS

HEARING ON CLAIMS-BASED DATABASES FOR POLICY DEVELOPMENT AND EVALUATION

Panel 1A—*Value, Purpose, Structure, Public Reporting and Policy Consideration*

Presented by:

Joel Slackman, Executive Director, Legislative and Regulatory Policy

Blue Cross Blue Shield Association

June 17, 2016

Good morning. My name is Joel Slackman, and I am an Executive Director for Legislative and Regulatory Policy for the Blue Cross Blue Shield Association (BCBSA). BCBSA is a national federation of 36 independent, community-based and locally operated Blue Cross and Blue Shield (BCBS) companies ("Plans") that collectively provide healthcare coverage for more than 107 million members – one in three Americans.

On behalf of BCBSA and its Plans, I would like to thank you for the opportunity to provide our perspective on all-payer claims databases (APCDs). For the following three reasons, we do not think now is the time for NCVHS to promote the current centralized APCD approach:

- First, centralizing data across multiple payers is expensive, time-consuming, and prone to issues of data integrity.

- Second, the recent Supreme Court's *Gobeille* decision greatly diminishes the value of a state's centralized data repository, and opens a Pandora's box of legal uncertainty about whether states can collect data from *any* ERISA (not only self-funded) plan.

- Third, states have a better way to go that is not fraught with the legal and regulatory uncertainties arising from the *Gobeille* decision: a distributed data approach. A distributed data infrastructure containing claims information on more than 30 million lives in the individual and small group markets is already operating in every state for CMS's risk adjustment program. States also have the opportunity to partner with entities like BCBSA to access their store of health care data for specified use cases.

Therefore, as NCVHS considers policy and reporting issues around claims-based databases, we urge you to consider the benefits and value of distributed data models, and the opportunities of partnering with the private sector.

TECHNICAL CHALLENGES TO CENTRALIZING DATA

BCBSA has considerable experience with the challenges in assembling meaningful claims across multiple, independent companies. Currently, BCBSA has a nationally recognized data initiative known as Axis, which exploits an enormous cache of claims resources for analytic purposes: more than 107 million people's worth of data, covering every zip code in the country, representing 93 percent of physicians and 96 percent of hospitals nationwide. BCBSA has relied on claims data from Blue Plans to provide insights into such issues as cost variations in knee and hip replace surgeries in markets across the U.S., the rapidly growing cost of specialty drugs, and differences in intensity of treatment between women and men 60 days following a heart attack.

It has taken us years of effort to achieve this ability to aggregate data in meaningful way. We are familiar with the inherent technical, security and operational challenges, and the need for sustained intellectual and financial capital, challenges that also face any state wishing to build an all-payer claims database (APCD). For example, the seemingly simple task of obtaining agreement on common data definitions across multiple claims systems to ensure apples-to-apples comparisons (e.g., is a mother and baby stay a combined admission or two separate admissions) was highly complex and required a sustained effort.

Maintaining the integrity and accuracy of centralized data is inherently challenging. Health care data is rife with often incompatible medical standards and coding schemes that require careful translation. Even an individual data provider's data is likely to come from many sources and be delivered in several tape or data formats. The industry's widely decentralized and largely autonomous data collection efforts make data quality a significant challenge.[1] When data are copied from multiple sources and then analyzed centrally, the burden of data processing falls on the analyzing institution, resulting in inefficient data cleaning. Because the only "source of truth" is the original source, there are large collections of errors (variant diagnoses, incompatible test results, and changes in coded data) that cannot be fixed without a deep understanding of the underlying source data.[2]

Even if centralized data are highly curated and extensively scrubbed – which we believe is not the case for all current state APCDs – issues come up where it is essential to call on the people who produced the data, which is difficult to do when the data are centralized. Mark McClellan made this point when he commented on CMS's original proposal to centralize data for risk adjustment: relative to a distributed model (which I discuss in more detail below), a third party's centralized model that aggregates data across multiple payers is at a disadvantage in assuring data consistency and quality, because the data are removed from those who know them best and are thus best suited to assess and address any idiosyncrasies or anomalies.[3]

LEGAL CHALLENGES TO CENTRALIZING DATA

The Supreme Court's decision in *Gobeille v. Liberty Mutual Insurance Company* – holding that the Employment Retirement Income Act of 1974 (ERISA) preempts a Vermont statute requiring the collection of health care data – raises significant issues about the value of investing in an APCD.

At issue in *Gobeille* was a self-funded group health plan sponsored by Liberty Mutual, which provided coverage to the company's Vermont-based employees and their dependents. The plan used Blue Cross

---

[1] Berndt, DJ., and others. (2001). "Healthcare Data Warehousing and Quality Assurance." *Computer*, v. 34, Issue 12.

[2] Carol C. Diamond, Farzad Mostashari and Clay Shirky. (2011). Collecting and sharing data for population health: A new paradigm. *Health Affairs,* v. 28 no. 2.

[3] Mark McClellan, "Comments on Proposed Rule, Standards Related to Reinsurance, Risk Corridors, and Risk Adjustment," October 31, 2011 (The Brookings Institution)

and Blue Shield of Massachusetts as a third-party administrator (TPA). Liberty Mutual ordered its TPA not to comply with the law, arguing that compliance would violate its own fiduciary duties; it then sued Vermont's implementing agency in federal court in Vermont, arguing that the Vermont APCD statute was preempted by ERISA and asking that the court enjoin enforcement of the statute.

As you know, the Court ruled against the plaintiff, Gobeille. In the short term, this ruling undercuts a major rationale for states investing in APCDs: as Gobeille argued to the Supreme Court, "An all-payer claims database that omits self-insured plans cannot be considered accurate or comprehensive,"[4] a point picked up by other briefs in favor of Gobeille:

- New York and 16 other states (plus DC) argued, "The usefulness of an APCD depends on its comprehensiveness. . . If self-funded plans are exempt from reporting their claims data to APCDs, the information in the resulting databases will be limited to Americans covered by traditional commercial insurance, Medicare, and Medicaid. That will skew the data in the APCDs toward the experiences of demographic groups more heavily represented by those forms of coverage. Furthermore, because Americans are likely to change insurance plans several times over the course of their lives, the omission of data from self-funded plans will make it difficult to track treatments and disease progressions over time."

- The Harvard Law School Center for Health Law and Policy Innovation wrote, "A database that included only claims from government payer programs, or even one that included claims from government payer programs and fully insured plans, would not be complete and therefore unable to approximate the health status of the American population as a whole. The population enrolled in self-funded insurance plans is a significant portion of health care users in this country and tends to be significantly younger and healthier than other populations. . . With sixty-one percent of the healthier commercially insured population potentially removed, APCDs will no longer provide an accurate portrait of the health of the general State population or of the enrollees of self-funded insurance plans."

Problematic as omitting self-funded claims is for APCDs, *Gobeille* could ultimately deprive APCDs of claims from *fully-insured* ERISA plans, virtually eliminating data on commercial populations. ERISA's "savings clause" at § 514(b)(2)(A) excepts from preemption any state laws that "regulat[e] insurance," and the language of the decision itself, on its face, appears to strike down the law as it relates to all ERISA-governed plans, whether insured or self-insured.[5] Indeed, we are aware that a number of health plans in states with APCDs have ceased submitting claims for any ERISA plans, self-funded and fully-insured. Future litigation is likely and casts a pall of uncertainty over any prospective APCD investment.

---

[4] http://sblog.s3.amazonaws.com/wp-content/uploads/2014/12/States-Reply-Brief.pdf

[5] Groom Law Group, "Gobeille v. Liberty Mutual Insurance Company: Supreme Court Ruling Calls into Question Validity of State APCD Statutes and Claims Taxes with Respect to ERISA Governed Plans," May 13, 2016.

The uncertainty extends to regulatory actions. For example, some have suggested that the Department of Labor could develop a national data solution. However, as noted in a March 2016 *Health Affairs* blog, this solution is harder than it sounds: no federal agency, whether the Department of Labor or HHS, currently collects anything like APCD, claim-level price, and quality data.[6] Even if one of these agencies agreed to collect plan data, to be effective it would have to be willing to gather the kind of timely, granular, and locality-specific data mandated by APCDs. Statistical or summary data would have little value to the type of analysis needed to assess, for example, whether the prices charged by a large health system jumped when they acquired a physician group. And, as Justice Ginsburg noted in her dissent, “It is unsettling . . . to leave the States dependent on a federal agency’s grace, i.e., the Department of Labor’s willingness to take on a chore divorced from ERISA’s objectives.”

DISTRIBUTED DATA MODEL AS A BETTER APPROACH

In general, so-called distributed or federated models provide a proven alternative to centralized data collection that can work around the legal issues and uncertainties arising from . A distributed network can perform essentially all the functions desired of a centralized database, while avoiding many disadvantages of centralized databases:

1. They allow data holders to maintain physical control over their data.

2. They ensure ongoing participation of individuals who are knowledgeable about the systems and practices that underlie each data holder’s data.

3. They allow data holders to assess and authorize query requests, or categories of requests, on a user-by-user or case-by-case basis.

4. Distributed systems minimize the need to disclose protected health information thus mitigating privacy concerns, many of which are regulated by the Privacy and Security Rules of the Health Insurance Portability and Accountability Act of 1996 (HIPAA).

5. Distributed systems minimize the need to disclose and lose control of proprietary data.

6. A distributed approach eliminates the need to create, secure, maintain, and manage access to a complex central data warehouse.

---

[6] Erin Fuse Brown and Jaime King, “The Consequences Of Gobeille v. Liberty Mutual For Health Care Cost Control,” March 10, 2016.

7. Finally, a distributed network also avoids the need to repeatedly transfer and pool data to maintain a current database, which is a costly undertaking each time updating is necessary.[7]

The FDA's Distributed Sentinel Model

In its amicus brief, the Harvard Law Center stressed the value of claims data by noting "Claims data allow researchers to study treatments once they are approved by regulatory bodies such as the Food and Drug Administration (FDA)." The FDA itself has adopted a distributed data model for such research. The FDA Amendment Act of 2007 required FDA to create the capability to use electronic health data from at least 100 million people to assess the safety of marketed medical products – the FDA responded by building today's distributed Sentinel System.

Sentinel now has more than 358 million person-years of data – including 4.0 billion prescriptions, 4.1 billion doctor or lab visits and hospital stays, and 42.0 million acute inpatient stays – which derive primarily from medical bills (claims), but a growing portion comes from EHRs or laboratory results (approximately 10 percent) – a portion expected to grow steadily in coming years.[8] Among its data partners are Anthem BCBS, BCBSMA, Kaiser Permanente, Humana, and Aetna. After creating a common data model, the FDA established a distributed analysis platform: users create and submit query (a computer program); data partners retrieve query; data partners review and run query against their local data; data partners review results; data partners return results via secure network.

The FDA has taken Sentinel beyond safety surveillance. For example, the FDA wanted to look at atrial fibrillation because it is a major public health problem (i.e., it increases risk of stroke, oral anticoagulation reduces risk, but many people for whom anticoagulation is recommended do not take it). FDA sent a rapid query to three data partners, looking at data for 16 million members over January 2006 to June 2014. The analysis found that 202,000 had AF plus additional risk factors, and 48 percent had no record of oral anticoagulant dispensing. This has led to the IMPACT-AF trial, a randomized trial under development to improve treatment with anticoagulants in people with AF.

CMS Risk Adjustment and EDGE Servers

A distributed data infrastructure containing claims information on more than 30 million lives in the individual and small group markets is already operating in every state.

---

[7] "Distributed Health Data Networks: A Practical and Preferred Approach to Multi-Institutional Evaluations of Comparative Effectiveness, Safety, and Quality of Care Jeffrey S. Brown, PhD, John H. Holmes, PhD, Kiran Shah, BA, Ken Hall, MDIV, Ross Lazarus, MBBS, MPH, and Richard Platt, MD, MSc, *Medical Care,* Volume 48, Number 6 Suppl 1, June 2010.

[8] "The FDA's Sentinel Initiative. The Food and Drug Administration has launched a new system to assess the safety of prescription drugs after they are approved for sale," *Health Affairs* (Health Policy Brief: June 4, 2015).

In deciding how to obtain and process claims data for ACA reinsurance and risk adjustment calculations, CMS rejected a centralized approach in favor of a distributed data collection model – CMS found that it proved the most effective approach because such a model would ensure minimal transfer of protected health information between issuers and CMS, thereby lowering privacy and data security risks; and standardization of business processes, timing and rules.[9]

CMS implemented the distributed data approach through External Data Gathering Environment or "EDGE" servers. Issuers upload enrollee, pharmaceutical claim, medical claim, and supplemental diagnosis information from their proprietary systems to an issuer-owned and controlled EDGE server. (Issuers have the option to own and operate the server themselves, or to have a third-party entity operate the server.) The EDGE server runs CMS-developed software designed to verify submitted data and execute the risk adjustment and reinsurance processes. Detailed data, file processing metrics, and outbound data files are provided to issuers, while only plan-summarized data and file processing metrics are provided to CMS. CMS does not receive any individual-level data as part of this process.

Edge servers are a valuable potential source of information because they contain data for both on and off exchange (except for grandfathered and grandmothered coverages, both of which are relatively a small percentage of these markets and diminishing). The Edge server contains medical (hospital and physician) and pharmacy claim data at a detailed level by de-identified member. The data stays under issuer control, but all the formats are standardized. True, Edge servers do not contain self-funded data, but they

OPPORTUNITIES FOR PARTNERSHIPS

Some briefs submitted in favor of Gobeille argued that APCDs needed self-funded claims data because "even if the majority of insurers cooperate to form a health care claims database, it will not deliver comprehensive data to researchers, limiting the usefulness of a private health care claims database." Moreover, "data from a single type of payer may only represent a fragmented portion of the American health care system."

We would take issue with the assertion that a private health care claims database cannot deliver the value of an APCD. As I noted in the beginning, BCBSA Axis has a vast repository of data that would deliver data equal in comprehensiveness to any current APCD. We are not keeping the data to ourselves, but entering into active partnerships with premier research institutions across the country to mine these data.

In that same vein, BCBSA could partner with State sponsored research initiatives through BCBS Axis Analytical Sandboxes, allowing for best of breed tools to be made available to mine the nation's largest store of commercial health care claims data. We are not talking about a plethora of studies as states with

---

[9] CMS, "March 31, 2016, HHS-Operated Risk Adjustment Methodology Meeting," March 24, 2016.

APCDs generate only a limited number of reports a year: in 2015 the state of Colorado, among the most prolific APCDs, published one study and began five others that involved commercial claims data.[10]

This approach, if focused on use cases of high value, would doubtless be more cost effective for states than building and maintaining a centralized APCD.

Distributed data models and public-private partnerships need not be mutually exclusive strategies: the former would support use cases that require not only a good population mix but also data across multiple payers; the latter, use cases for which deep data across all lines of business would be most beneficial.

That concludes my testimony. Thank you again for the opportunity to present BCBSA's perspective to the Committee.

---

[10] "Cost Driver Spot Analysis: Avoidable ED Use," (November 2015). Studies in progress: "Impact Study of Project Angel Heart's Home-Delivered Meals,"; "Cost Transparency in Colorado Hospitals,"; "Identify Opportunities to Reduce Use of Potentially Harmful Medications During and Post Surgery,"; "Assessing the Impact of Cost-Sharing Policy Changes on the Use of Preventive Services,"; "Analysis of Inflammatory Bowel Disease Treatments and Medications to Identify Best Practices."