CONFERENCE REPORT

NAHDO 2024

4th Annual Meeting and Networking Sessions

Telling the Story:
Strategies for Improving Healthcare Data

May 7-8, 2024
Hotel Monteleone
New Orleans, Louisiana

Version 1.0

EXECUTIVE SUMMARIES

APCD Development Forum

This session focused on sharing learnings and best practices between different states involved in developing and implementing All-Payer Claims Databases (APCDs). Here are the key takeaways:

- **Communication**
  - Importance of having public-facing tools (e.g., dashboards) to showcase data and gain user interest.
  - Value of clear communication plans with designated liaisons to address user questions and minimize data request challenges.
  - Targeting users with specific data use cases in mind.
- **Data Processing**
  - Benefits and considerations of outsourcing data processing.
  - Importance of clear and user-friendly data documentation.
  - Maintaining good relationships with data submitters.
  - Keeping data processing methods simple and versioning easy to understand.
- **Policy and Governance**
  - Finding legislative champions to advocate for APCDs and secure funding.
  - Using APCD data reports to inform policy decisions at the district level.
  - Highlighting the value proposition of APCDs for policymakers (e.g., revenue generation through research partnerships).
  - Possible action item: Sharing data usage agreements (e.g., multi-use agreements) as successful models.
- **Analytic Methods**
  - Standardizing data request processes (e.g., tiered access levels) to streamline workflows.
  - Utilizing tools like pre-formatted data sets and sprint planning approaches to improve efficiency.

The session also highlighted the challenges of navigating different state regulations and data privacy restrictions. There is a need for collaboration and knowledge sharing between states to develop best practices for data processing, communication, and policy development.

Data Quality Forum

The NAHDO Data Quality Forum provided a platform for attendees to discuss data quality challenges and best practices. Julia Tremaroli led the session, offering an overview of the forum's purpose and facilitating updates on key topics: the Non-Claims Payment Layout (NCP) and race/ethnicity data collection in Medicare.

Janice Bourgualt presented the latest NCP Layout, a standardized format for non-claims payment data exchange (available [here](#)).

Heather Koenig from the Office of Inspector General (OIG) addressed limitations in Medicare's race and ethnicity data collection. She highlighted a relevant OIG study and a resource guide for using this data effectively. A key takeaway was the importance of prioritizing self-reported race and ethnicity data whenever possible.

The forum also served as a springboard for attendee engagement. Discussions centered on various data quality concerns, including:

- The use of "other" codes in place of specific data.
- The appropriate level for addressing missing race/ethnicity data (dashboard, submitter, or both).
- Techniques for quickly assessing data quality changes between extracts.
- The significance of understanding data collection methods and documenting data history (metadata).
- Exploring alternative local sources for race/ethnicity data beyond CMS sources.
- Implementing Data Quality Atlases, similar to the MHDO model, to visualize data quality for researchers and garner policy support for addressing data issues.
- Recognizing that data quality challenges may vary based on data submitter size, particularly during initial data collection periods.
- Sharing best practices for engaging providers and submitters through forums or public comment periods, as seen in states like Texas and New Hampshire.
- Balancing data quality resources with needs – how much is enough for effective QA?
- Communicating data quality concerns constructively, without compromising the data set's value.
- Utilizing data flagging for questionable elements, and considering data release implications.
- Identifying relevant metrics for measuring and managing data quality. The discussion acknowledged that all data contains errors, and data quality assessment should be use-case specific.
- Prioritizing timeliness versus completeness, depending on the specific data field and intended use.
- Acknowledging that data use itself can improve data quality.
- Emphasizing the importance of a review period before publishing or releasing data analysis for public consumption. This allows for the identification and correction of potential data issues.

Session 1A. Identifying Data Gaps and New Data Sources that Can Help Address Them

This NAHDO session focused on identifying data gaps and exploring potential solutions. Leanne Candura and Kathy Hines led the discussion, facilitating brainstorming and group work around key themes.

Critical Data Gaps:

- **Lab Results:** Absence of lab data limits clinical insights. Potential solutions include hospital data collection or supplemental file submissions, acknowledging limitations in population coverage. Existing successful models from other states could be shared.
- **Social Determinants of Health (SDOH):** Missing race/ethnicity and SOGI data pose challenges. Backfilling race/ethnicity data for Medicaid enrollees using enrollment data and sharing discharge data were discussed as possibilities. Additionally, exploring Z codes (administrative codes for social factors) and provider training to improve data capture were considered. Utilizing Social Security data for race/ethnicity was deemed unreliable due to data quality issues. Legislation mandating race and ethnicity data collection might be necessary.
- **Vital Records:** Linking with vital records requires collaboration with other data agencies, but legal restrictions in some states prevent this. Universal hashing with Master Patient Index (MPI) could facilitate linking when permitted. However, de-identification practices could hinder this approach. Policy changes or legislation might be needed to allow linking vital records data with APCD or discharge data.
- **Workforce and Provider Data:** Gaps could be filled through surveys or utilizing existing federal data sources like NPPES, PECOS, and hospital system compendiums. Leveraging licensure data was also explored, but limitations exist due to some providers being unlicensed or claims being billed under doctors (particularly in behavioral health). A stakeholder group could help identify and address these issues.
- **Pharmacy Benefit Manager (PBM) Data:** While Arkansas recently passed legislation to collect PBM information, it was challenged by PBMs. Pharmacist associations support PBM data collection due to its potential impact on tracking medication costs. Challenges include obtaining PBM rebate data at the drug level and overcoming PBM lobbying efforts against data sharing. NAHDO might be able to assist PBMs in understanding the inapplicability of antitrust concerns in this context (referencing [CIVHC Antitrust summary](#)).
- **Health Encounter Data Not Billed to Payer:** EMS and community response data often fall outside the traditional payer billing system, creating gaps. Efforts are underway in some areas to integrate this data, but consistency and legal restrictions remain hurdles. Bi-directional data exchange between patient and population levels with appropriate data aggregation and disaggregation could be a potential solution.

By identifying these data gaps and exploring solutions, NAHDO members can work towards a more comprehensive data landscape for informed healthcare decision-making.

Session 1B. Communicating the Value of Data

This NAHDO session focused on successful data dissemination strategies, emphasizing clear communication and audience engagement. Key takeaways included:

- **Data Dissemination as a Foundation:** Nevada's experience highlighted the power of data dashboards, particularly during COVID-19. Their modular design, standardized reporting structure, and accessibility features facilitated effective communication.
- **Targeted Communication Strategies:** Defining the audience is crucial. Different approaches are needed for students, researchers, policymakers, and legislators. Marketing strategies should explain APCD's purpose, use cases, security measures, and the value proposition for each audience.
- **Creating Engaging Dashboards:** Understanding user needs and learning styles is essential. Dashboards should be clear, concise, visually appealing, and frequently updated.
- **Balancing Usefulness and Engagement:** Effective dashboards are both informative and engaging. Clarity, actionable insights, and a focus on user needs should guide dashboard design.
- **Mission-Driven Communication:** Data dissemination should prioritize real-time accuracy, transparency, and balancing the needs of diverse users. These principles should inform the design and content of data dashboards.
- **Engaging Users:** Interactive dashboards with accessibility considerations enhance user experience. Cater content to different education levels and utilize relatable comparisons to foster understanding.
- **Meeting Users Where They Are:** Utilizing common and standardized data variables builds trust and facilitates user comprehension. Engaging stakeholders helps ensure data integrity and a unified voice. Social media can be a valuable tool to reach different audiences and promote data availability.
- **Responding to Current Events:** Data communication should be agile and responsive to current events. Understanding the limitations of historical data and identifying alternative data streams are crucial for informed decision-making.
- **De-emphasizing Tools:** While tools like Tableau and Power BI play a role, the focus should be on defining use cases and user needs. Tools should be selected based on their ability to address specific data intake, quality, documentation, and delivery needs. Security considerations, balancing data accessibility with security, and securing buy-in from stakeholders are all important aspects.
- **Synthetic Control Methods:** Kentucky's use of VA augmented synthetic control methods to assess the impact of Managed Care Organizations (MCOs) demonstrates the potential for innovative data analysis techniques.

By implementing these communication strategies, NAHDO members can ensure their data resources are effectively disseminated and utilized by a wide range of stakeholders.

Session 2A. Considerations for Data Linkage

This NAHDO session explored key considerations for data linkage, a process of combining data sets from different sources to gain richer insights. The discussion focused on three key themes: privacy and security, accuracy and reliability, and identifying appropriate data sources.

Privacy and Security:

- Balancing the value of research questions (e.g., linking claims data with air quality data) against privacy concerns is a challenge. Potential solutions include enclave models for data storage or third-party data summarization to protect confidentiality.
- Regularly testing data security systems through simulated attacks was identified as a best practice for identifying and addressing vulnerabilities.
- While public use files offer privacy protections, their usefulness can be limited due to computational complexity. Striking a balance between privacy and analytic utility is crucial.

Accuracy and Reliability:

- Backfilling race/ethnicity data using census data proved successful in Arkansas, but limitations were identified in areas with a high prevalence of ERISA self-funded plans. This highlights the need for comprehensive data gap analysis.
- Louisiana utilizes probabilistic matching software (even the free version) to link vital records, COVID data, and crash data with an 85% success rate.
- Nevada's success story involved linking public health and Medicaid data sets to create a comprehensive master data set (e.g., for syphilis) that was securely shared with authorized agencies.

Finding the Right Data Sources:

- The importance of vendor expertise was emphasized. A cautionary tale highlighted the risks of selecting vendors with limited understanding of the data sets being linked.
- Massachusetts linked disparate data sets (cancer registry, courts, jail data) to create a de-identified master health data set. They collaborated with Public Health to assign unique identifiers and facilitate data exploration across various domains.
- Another example underscored the importance of data review processes. A vendor error in coding a disposition field (dead vs. discharged to nursing home) was caught during provider review, preventing the publication of inaccurate data. This highlights the value of multi-layered data review before dissemination.

By carefully considering these aspects, NAHDO members can leverage data linkage to enhance data analysis and inform healthcare decision-making, while ensuring responsible data stewardship.

Session 2B. Identifying the Value of Your Data Resources

This NAHDO session aimed to empower participants to extract maximum value from their data resources. Through interactive exercises, attendees explored three key areas:

1. **Identifying Hidden Value:** The session kicked off by guiding participants to uncover the hidden potential within their data. Techniques like data mining and correlation analysis were presented as tools to discover new products and services that data can support.
2. **Crafting Compelling Narratives:** Developing clear, concise presentations that showcase the impact of data was a key focus. The session emphasized translating data insights into actionable solutions for potential users. Highlighting success stories and real-world examples was identified as essential for building a strong narrative.
3. **Collaboration for Innovation:** Sharing best practices and brainstorming with peers was championed as a powerful driver of innovation. The session fostered an environment where participants could learn from each other and develop improved approaches to data analysis.

A core concept explored throughout the session was the "data analytics continuum," which categorizes analytics by complexity:

- **Descriptive Analytics (foundational):** This level provides foundational insights through reports, dashboards, and historical data analysis. It's crucial for understanding past events and informing other analytics. Key takeaways included ensuring a clearly defined question to guide analysis, identifying and transforming data to answer the question effectively, and leveraging simple descriptive tools to make data accessible for users. This foundation is essential for understanding data customers and their needs.
- **Explanatory Analytics:** This level delves deeper to understand "why" things happened. Techniques like data mining and correlation analysis were explored. An example application discussed was a new product designed to help employers analyze healthcare cost drivers. Key considerations for explanatory analytics included developing a scalable, cost-effective tool for smaller employers and identifying the necessary data to provide employer-specific insights while building trust in its accuracy.
- **Predictive Analytics:** This level uses statistical modeling to forecast future events. The session highlighted the importance of identifying target audiences and the specific problems that predictive analytics can solve for each group. Examples included using data to manage health during COVID-19 (vaccine distribution, herd immunity) and price transparency for consumers. Communicating the limitations and volatility of predictions to users was emphasized, along with building trust through accurate predictions and successful demonstrations.
- **Prescriptive Analytics (most complex):** This level provides recommendations for achieving desired outcomes, using simulations and decision support tools. Examples included Kentucky Medicaid's use of synthetic controls to assess the impact of MCO models on spending and the Accorded Model, which evaluates cost implications of interventions. A key takeaway was the potential of prescriptive analytics to inform

strategies, price negotiations, and vendor evaluations through real-time insights into intervention effectiveness.

The session also addressed additional topics:

- **Marketing Your Analytics:** Strategies for promoting data analysis tools were explored, emphasizing user needs and data privacy considerations.
- **Promoting the Value of Your Analytics:** The potential of data sales to the commercial sector was highlighted, showcasing its positive impact and safeguards. Stakeholder engagement and continuous innovation were identified as key to success in this area.

By following these insights, NAHDO members can transform their data into a powerful asset for informed decision-making.

Session 3A. Data Sharing

This session at NAHDO emphasized the importance of trust in data sharing initiatives. Attendees discussed concerns like data quality, potential misuse, security breaches, and unclear ownership. Strategies to address these concerns involve data quality measures, clear Data Use Agreements (DUAs), robust security protocols, and transparent communication about data protection and value proposition. Challenges like balancing audit frequency with data volume and limited penalties for misuse were acknowledged.

The session also explored technical challenges to data sharing. These challenges include a lack of standardization across data formats, security concerns, and the complexities of varying data privacy laws and agreements. Possible solutions involve adopting cloud-based solutions and common data dictionaries, investing in robust security measures, and streamlining data privacy laws with standardized DUA/DSA templates. User education and mentorship programs for data professionals can also play a critical role.

Effective communication is essential for successful data sharing efforts. Communication should address data updates, data quality limitations, and the legal aspects of data access and usage. It should also clearly articulate the value proposition for data providers and users, outlining the processes involved in data collection, redistribution, and issue resolution. Technical details regarding data formats, security measures, and access controls should also be communicated.

Data users can expect regular updates through reports, dashboards, and presentations. Data providers should be kept informed about how their data is being used and the value it contributes. Regular communication with funders and policymakers should highlight the impact of data sharing on decision-making processes and funding outcomes. Standardized communication practices, stakeholder education, and clear communication regarding legal aspects are all important for ensuring effective communication.

The future of data sharing holds the promise of standardized data formats, a national API registry, and a more cost-effective data sharing infrastructure. Enhanced security measures, improved data quality scores, and expanded control over data privacy through Do Not Share (DNC) rules are also envisioned. Real-time data updates, a centralized data repository, and automated data processing workflows can streamline data sharing and improve accessibility. Increased data literacy and education for all stakeholders, along with a central location for data sharing best practices, will be crucial for transparency.

The ultimate goal is a sustainable data sharing ecosystem with uniform data policies, shared resources, and enforced data quality requirements. This will lead to improved care coordination, faster research, reduced administrative burdens for providers, and a more empowered patient population through secure, patient-controlled data access. Challenges include balancing standardization with the need for innovation, addressing privacy concerns, achieving stakeholder buy-in, and implementing a standardized data sharing infrastructure.

Session 3B. The Return of the Patient Centric Data Model

Dawniece Trumbo, Sara Hallvik, Katie Cadigan

During this session, we explored what it means to ground ourselves in the patient experience as we design and implement data models to improve the health outcomes and health care systems in our communities.

After hearing two very different versions of "Marcella's" story, one in which systems use data cohesively and efficiently to improve her care, and the other in which data systems are disjointed and fail her, we reflected on our own experiences as patients. We identified ways that data could be better leveraged to improve our experiences in the healthcare system, such as optimizing data flows to improve provider to provider communication so that patients do not need to repeat their stories. An important theme was around the need to balance standardization with personalization to maximize efficiency while remaining adaptable to the specific needs of individuals.

From there, we engaged in small group discussions to brainstorm specific ways that data sources could be identified, coordinated, and analyzed to support patients.
Each group worked with a unique patient profile around the same scenario of the patient with high blood pressure being prescribed a new medication. We came up with creative ideas as well as questions to consider and explore. For example:

- Robert, a veteran living alone on a fixed income: How can VA and non-VA claims be combined to create a more holistic picture of needs and experiences?
- Pavi, a single parent and recent immigrant whose primary language is Urdu: community health workers and other community-based support could be key supports; how can data inform their outreach and intervention efforts?
- Sam, who is unhoused and uninsured and uses substances: what feedback lops could be put in place to best understand experiences across multiple systems?
- Denise, a 40 yr. old African American woman who just delivered her first baby: longitudinal data could help identify new care caps that emerge after delivery.
- Taylor, a recent HS graduate who moved to Hollywood to be a movie star and is currently bartending: a hub and spoke model could help provide stability and consistency while residence and employment might be variable.
- Alex, a non-binary accountant working remotely from Alaska: telehealth could be a modality to increase engagement.

Some overarching themes from these discussions include:

- What assumptions, i.e., about gender, might we be making about our patients given the limited information we had. And how do those assumptions impact choices we make about data flows and analyses?
- For some of our patients, broader social determinants of health, i.e., housing, may need to be addressed before focusing on medication adherence.
- Who might be missing or hard to find in data sources, and how can we address that?

Session 4A. Opportunities to Optimize APCD Operations

This summary explores the opportunities and challenges faced by All Payer Claims Databases (APCD) in optimizing their operations.

Funding

APCDs currently face challenges with unreliable grant funding, limited revenue streams from membership fees and data sales, and staffing constraints that hinder fundraising efforts. The loss of a major funder has further exacerbated the situation. To ensure long-term sustainability, APCDs need a diversified funding strategy. Potential solutions include tiered subscription models for data access, increased membership fees, targeted marketing, data sales to other federal agencies, and improved communication showcasing the value proposition of APCD data. Developing a transparent pricing structure and long-term planning documents will also be crucial.

Data Analytics

Data quality issues pose a significant challenge for APCD analytics. These issues include data inconsistencies, duplicate entries, and potential for hidden problems in submitted data. Limited analytical capabilities and workflow inefficiencies further hinder efforts. Solutions involve implementing cloud-based versioning for better data tracking, enhanced validation processes, and exploring advanced analytics tools. Streamlining workflows by involving analysts earlier and fostering better communication with data submitters will also improve data quality. Prioritizing automation of data cleaning and validation tasks, along with stakeholder feedback mechanisms, will further enhance the analytical process.

Data Management and Quality

APCDs can prioritize data quality through clear and user-friendly data submission guidelines, regular stakeholder feedback opportunities, and data validation checks throughout the processing pipeline. Benchmarking against data from state agencies and assigning data review tasks to qualified personnel are additional strategies. Early integration of state data files, flexible data processing to adapt to changes, and consistent data review schedules are also important practices. The Data Quality Unit, Payer Management Team, Data Processing Team, and Analytics Team all play crucial roles in ensuring data quality.

Data Privacy, Security, and Governance

(1) **Use of A.I**. APCDs recognize the importance of safety and ethical use of AI in data analysis, along with the need for clear use cases and human oversight to mitigate bias and security risks.
(2) **Data Governance Practices** Good data governance requires a standardized data request process and clear data sharing protocols (with statistical de-identification techniques). Two challenges in governance are the need for in-house expertise and

addressing inconsistencies across state models. Education and training for staff on data privacy and security best practices, along with staying informed about evolving data privacy regulations, are essential considerations.

Client Engagement and Management

Many APCD clients lack the data expertise to fully utilize available resources. Tiered data products, user training programs, and improved data quality documentation can address this challenge. Establishing clear communication channels, utilizing Data Use Agreements, and implementing a data request application with a transparent queue will improve communication and manage client expectations. Ongoing client engagement strategies are also important for user satisfaction and continued data utilization.

APCDs for Policy and Data Efficiency

(1) Bridging the gap between policymakers and APCD leadership is critical for ensuring data-driven policy decisions. It would be helpful to have a stakeholder committee that includes policymakers and data users to increase communication and inform policy development. Investing in data literacy initiatives for policymakers and fostering a collaborative environment are also important considerations.
(2) Periodic reviews of data governance policies and encouraging flexible frameworks will ensure data resources are effectively utilized.

Conclusion

By addressing these challenges and implementing the proposed solutions, APCDs can build a more sustainable and efficient organization that empowers users to leverage the full potential of its valuable data resources for public health research, analytics, and informed policy decisions.

Session 4B. What's New in Analytics

This summary explores four key trends shaping the field of data analytics: Social Determinants of Health (SDOH), Linking Big Data Sets, the rise of Artificial Intelligence (AI), and the potential of Data Lakes.

1. Social Determinants of Health (SDOH):

APCDs are actively collecting and analyzing SDOH data to understand the social factors influencing health outcomes. Data sources include the Social Vulnerability Index (SVI), the Bayesian Surname Model (BISG), and the Area Deprivation Index (ADI). Efforts are also underway to link hospital discharge data with Z codes (social factors contributing to health conditions) and Sexual Orientation and Gender Identity (SOGI) data. Challenges remain in ensuring the granularity of geographic data and the consistency of SDOH data collection across healthcare providers.

2. Linking Big Data Sets:

The ability to link disparate data sets offers a more holistic view of various complex issues, including SDOH. However, this process presents challenges in data governance, security, access controls, and data cleaning to ensure alignment. Despite these hurdles, APCD is exploring linkage opportunities with data from cancer registries, birth and death records, and provider data sources.

3. Artificial Intelligence (AI):

AI is making significant strides in the field of data analytics. APCD is exploring the use of AI tools like Python, R, SAS, STATA, and AI for various tasks, including automation, fraud detection, pattern identification, coding practice analysis, and data translation. These advancements hold promise for improving data quality and efficiency. However, concerns remain regarding potential biases in models trained on existing data, highlighting the importance of transparency in building trust.

4. Data Lakes:

Data lakes offer a centralized repository for compiling data from numerous sources simultaneously. This approach has the potential to unlock valuable insights into SDOH and other areas of public health research. However, careful consideration of use cases and navigating data governance regulations across various sources are essential challenges that need to be addressed.

By actively exploring these four key trends, APCD is well-positioned to leverage the power of data analytics for a more comprehensive understanding of public health issues and to ultimately improve health outcomes for all.

Session 5A. Governance is Critical to a Successful Program

Facilitators - Angela Taylor and Craig Schneider

This session underscored the importance of data governance, the framework that ensures the quality, security, and accessibility of data, for successful programs. The discussion centered on several key areas:

- **Building and Maintaining Data Governance Programs:** HDOs need clear steps to establish a data governance program. Ongoing tasks for established programs include managing data access and adapting to evolving regulations.

- **Curating and Maintaining Data:** Finding best practices and data sources from other institutions is crucial. Careful consideration must be given to legal and ethical concerns when determining what data can be released, to whom, and under what conditions. Following up on data usage and adapting release policies to a changing landscape are also important aspects.

- **Balancing Inclusivity and Efficiency:** Striking a balance between providing broad access to data and ensuring efficient processing of requests is a challenge. Standardized agreements and tiered access systems can help. Prioritization becomes important, and cost structures can be designed to be inclusive of different user types.

- **Evolution and Adaptation:** Data governance needs to be adaptable to keep pace with new trends, technologies, and regulations. Recent years have seen increased security protocols and legal scrutiny as examples of such adaptations. Building flexibility into the governance structure from the beginning allows for easier adjustments in the long term.

- **Sharing Experiences:** Participants discussed successful approaches to data governance, such as clear documentation and strong stakeholder buy-in. Bottlenecks in the data release process, ensuring timely access, and maintaining data quality were identified as challenges. Sharing best practices for data governance documentation was seen as an important way to move forward.

By following these key principles and fostering collaboration, organizations can establish effective data governance programs. These programs ensure the responsible use of data, which is critical to the success of any program that relies on data-driven decision-making.

Session 5B. Multistate Projects

This summary explores the potential and challenges of multistate data sharing projects for public health research.

Use Cases and Opportunities

Several use cases highlight the value of collaboration:

- **Fentanyl Crisis:** A data sharing partnership between Oregon and Washington could track prescribing patterns, treatment data, and overdose locations across state lines, informing policy and improving treatment access.
- **Workforce Tracking:** Analyzing data from multiple states can help identify healthcare providers serving specific populations, particularly in rural areas with limited resources. This is critical for workforce planning and resource allocation in behavioral health and primary care.
- **Maternal and Child Health:** Tracking service utilization across state lines, especially for vulnerable populations, can improve public health responses and outcomes.
- **Regional Analysis:** Understanding why patients seek care outside their home state, as seen in New England for specialized services, can inform healthcare delivery and resource allocation across regions.
- **Disaster Preparedness:** Sharing patient data across state lines during emergencies like hurricanes or wildfires can improve response efforts and resource allocation. This includes planning for patient surges and ensuring insurance portability.
- **Disease Surveillance:** Collaborative disease registries can track population movement and improve public health surveillance, especially in border regions. This can help identify centers of excellence and potential outbreaks.
- **Antitrust Analysis:** Multistate data analysis can be used to assess the impact of healthcare mergers and acquisitions across markets, helping to identify potential anti-competitive practices.

Challenges and Considerations

Despite the benefits, multistate projects face challenges:

- **Data Tracking:** Tracking individuals moving across state lines requires robust data sharing agreements and standardized data collection methods.
- **Data Governance:** Clear agreements regarding data access, security, and privacy are crucial for collaboration.
- **Data Quality:** Ensuring consistent data formats and implementing robust quality control processes are essential for reliable analysis.
- **Stakeholder Engagement:** Effective communication and collaboration between participating states and public health agencies are key to project success.

Moving Forward

By addressing these challenges and leveraging the power of multistate data collaboration, public health agencies can gain valuable insights to improve public health policy, resource allocation, and overall population health outcomes.

LIST OF CONFERENCE ATTENDEES

Bola Adams	Population Health Analyst, Human Services Research Institute
Olga Armah	Manager of Research and Planning, CT Office of Health Strategy
Meredith Arrison	Texas Department of State Health Services
Jason Aziz	Director of Health Economics, New Hampshire Insurance Department
Josiane Bechara	Senior Research Methodologist, NORC at the University of Chicago
Monica Begay	Director Data Analytics & Quality, SYNCRONYS
Alexia Benshoof	Health Bureau Chief, Nevada Department of Health and Human Services (Office of Analytics)
Tanya Bernstein	Executive Vice President, Freedman HealthCare
Pragya Bhattarai	Medical Research Specialist, Texas Department of State Health Services
Kristina Bondurant	Director and Epidemiologist, Data Sciences and Public Health Programs, Arkansas Foundation for Medical Care
Erin Bonney	Director, Health Informatics and Reporting, Center for Health Information & Analysis (MA)
Srimoyee Bose	Director of Research, One Utah Health Collaborative
Janice Bourgault	Senior Consultant, Freedman HealthCare
Brad Brockbank	Sr. Manager, Healthcare Data Markets, Peterson Center on Healthcare
Tarik Brown	Director, Texas Department of State Health Services - THCIC Unit
Katie Cadigan	Manager, System-wide Quality Improvement, Comagine Health
Leanne Candura	Vice President, Human Services Research Institute
Jason Caplan	Chief of Special Projects, Maryland Health Care Commission
Mike Carson	Lead Architect, hMetrix
Lesia Carter	Assistant Director/Grant Specialist, Arkansas Insurance Department
Colin Caruso	Account Executive, Milliman MedInsight
Sandra Chao	Senior Researcher, Mathematica
Frank Cheung	Co-Founder & CEO, Accorded
Alansia Cicero	Data Reporting Analyst, Virginia Health Information
Andrea Clark	Director, Health Care Economics, CareQuest Institute for Oral Health
Renee Clark	Manager, Care Coordination and Interoperability, CareQuest Institute for Oral Health
Amy Costello	University of New Hampshire
Scott Curley	Manager, Privacy & Compliance, Center for Health Information & Analysis (MA)
Megan Denham	Senior Research Associate, Georgia Tech Research Institute
Maushami DeSoto	Health Research Services Administrator, Agency for Healthcare Research and Quality
Jesse Drummond	Director of Data Analytics, Onpoint Health Data
Jon Duke	Principal Research Scientist, Georgia Tech Research Institute
Joanna Duncan	COO, Onpoint Health Data
Lisa Dunning	BI System Data Analyst, MMCAP Infuse
Matt Enright	Director of Data Analytics, Delaware Health Information Network
Mary Ann Evans	Research Analyst, Oregon Health Authority
Nicole Fawcett	Director of Operations, Virginia Health Information
Karl Fernstrom	Health Care Data Service Center Manager, Minnesota Department of Health
Michael Fields	Sr. Principal Program Manager, UnitedHealthcare
Helen Figge	CSO, MedicaSoft
Kimberley Fox	Senior Research Associate, University of Southern Maine and Maine Health Data Organization
Cari Frank	VP | Communication and Marketing, Center For Improving Value In Health Care
John Freedman	CEO, Freedman HealthCare
Hannah Friedman-Bell	Advisory Services Analyst, Mathematica
Rik Ganguly	Data Science Consultant, Freedman HealthCare
Sule Gerovich	Senior Fellow, Mathematica
Lorie Geryk	Senior Program Manager, WA Health Care Authority
Stefan Gildemeister	Health Economics Program Director, Minnesota Department of Health
Venkateswar Gopisetti	VP - Strategic Growth and Alliances, Datagaps Inc.
Lynn Goyne	Vice-President, System13, Inc.
Eric Guroff	CEO, Ratio PBC
Sara Hallvik	VP, Data Solutions, Comagine Health
Morgan Harris	Director of Healthcare Analytics, New Hampshire Insurance Department
James Harrison	President/CEO, Onpoint Health Data
Joseph Harrison	Data Architect, UTHealth Houston's School of Public Health
Nerissa Harvey	Assistant Manager HDDS, Tennessee Department of Health
Nathan Hedberg	Director, Product Strategy, hMetrix
William Hendon	Project Manager, Freedman HealthCare
Kathy Hines	Senior Health Care Data Advisor, Center for Health Information & Analysis (MA)
Elizabeth Holcomb	Director, Georgia Office of Health Strategy and Coordination
Darcy Holladay Ford	Director of Research, Center For Improving Value In Health Care

Cameron Holstead	PhD Economist, Research & Planning Consultants	Preston Morris Jr	President/CEO, System13, Inc.
Jennifer Howley	UHFDDS Data Scientist, NH DHHS Division of Public Health Services	Maggie Mueller	Director of Data Operations, Center For Improving Value In Health Care
Oscar Ibarra	Associate Director, State of Maryland - HSCRC	Kara Nester	Researcher, Mathematica
Neetu Jain	Researcher, Mathematica	Kate O'Neill	Director of Health Systems Data and Analytics, Green Mountain Care Board
Don Jang	VP, NORC at the University of Chicago	Amshe Ogun	Epidemiologist II, Tennessee Department of Health
Will Johnson	Freedman HealthCare	Pam Owens	Director, Center of Healthcare Data and Analytics, Agency for Healthcare Research and Quality
Sely-Ann Johnson	Project Manager I, Freedman HealthCare		
Richard Jones	Senior Vice President, Advanta Government Services, LLC	Evaren Page	Director of Science and IRB, Oklahoma State Department of Health
Generosa Kakoti	Director, Office of Healthcare Statistics, Tennessee Department of Health	Twanisha Parnell	Chief Financial and Administrative Officer, Center For Improving Value In Health Care
Amy Kinner	Director of Health Analytics, Onpoint Health Data	Justine Pascual	Research Analyst, SEIU United Healthcare Workers - West (SEIU-UHW)
Heather Koenig	Social Science Analyst, U.S. Department of Health and Human Services, Office of Inspector General	Kristin Paulson	President and Chief Executive Officer, Center For Improving Value In Health Care
Güneş Koru	Professor, University of Arkansas for Medical Sciences	Anne Paumgarten	Health Care Data Analyst, Green Mountain Care Board
Chris Krawczyk	Chief Analytics Officer PhD, HCAI - Office of Information Services (CA)	Laurie Paxson	Data Architect, Commonwealth of Kentucky
		Lauren Peters	Executive Director, Center for Health Information & Analysis (MA)
Mike Latterner	Sr Data Scientist, NORC at the University of Chicago	Becca Peterson	APCD Client Services Specialist , Virginia Health Information
Jeffrey Leintz	Vice President and Director, NORC at the University of Chicago	Donald Poe	Technical Director, Arkansas Center for Health Improvement
Lisa Li	Staff Epidemiologist, Oklahoma State Department of Health	Jo Porter	APCD Council Co-Chair, University of New Hampshire
Sarah Lindberg	Senior Data Science Consultant, Freedman HealthCare	Erin Proven	Business Intelligence Specialist, Louisiana Department of Health, Office of Public Health
Brad Long	Senior Health & Human Services Principal, Resultant	Victoria Razin	Research Engineer, Georgia Tech Research Institute
Jenny MacKenzie	Policy Associate I , University of Southern Maine and Maine Health Data Organization	Dana Richardson	CEO, Wisconsin Health Information Organization
Robert MacLean	General Counsel, Center for Health Information & Analysis (MA)	Jillian Rider	Lead Analyst, Virginia Health Information
Katie Martin	President and CEO, Health Care Cost Institute	Mel Riffe	Lead Programmer/Analyst, System13, Inc.
Britteny Matero	Partner and Senior Vice President, Innsena	Gina Robertson	Onpoint Health Data
Jonathan Mathieu	Senior Consultant, Freedman HealthCare	Arneris Rojas	Project Associate , Freedman HealthCare
Kevin McAvey	Managing Director, Manatt Health	Kyle Russell	CEO, Virginia Health Information
Paul McCormick	VP of Data Operation, Center For Improving Value In Health Care	Seema Saraswathy	Research Specialist, Texas Department of State Health Services
Paul Messino	Director, State Health, Mathematica	Lori Savoie	Program Director, Utah Department of Health & Human Services
Joseph Miller	Office Director of Public Health Statistics, Mississippi State Department of Health	Craig Schneider	Principal, HMA
Pam Mink	Health Services Research Unit Supervisor, Minnesota Department of Health	Lauren Sears	Manager, Manatt Health
		Yujing Shen	Health Economist, RELI Group
Kenley Money	Director of Information Systems Architecture, Arkansas Center for Health Improvement	Suzanne Shumway	Meeting Magician, University of New Hampshire
Liz Mooney	VP	Research, Partnership & Innovation, Center For Improving Value In Health Care	Susan Slappey	Program Operations Administrator, Agency for Healthcare Administration (FL)
Anelia Moore	Assistant Director and Senior Policy Advisor, Georgia Office of Health Strategy and Coordination	Marissa Smith	Project Manager, Freedman HealthCare
		Lee Spangler	Assoc. Professor & Exec. Dir. , UTHealth Center for Health Care Data
Kyra Morgan	State Biostatistician, Nevada Department of Health and Human Services (Office of Analytics)	Audrey Speter	Director, hMetrix
		Colin Stauffer	Director of Data and Development Services, Resultant

Megan Stead	Senior Research Director II, NORC at the University of Chicago
Ben Steffen	Executive Director, Maryland Health Care Commission
Andrea Strong	Chief - Financial Data Administration, State of Maryland - HSCRC
Angela Taylor	Biomed Informatics Data Architect, Commonwealth of Kentucky
Chang Thao	Data Analyst, MMCAP Infuse
Jen Thompson	Health Program Manager, Nevada Department of Health and Human Services
Jenn Toms	Director of Business Development, Onpoint Health Data
Julia Tremaroli	Manager, Program Operations, Integrated Healthcare Association
Huong Trieu	Senior Director of Research , Center for Health Information & Analysis (MA)
Dawniece Trumbo	Director of Analytics, Virginia Health Information
Alex Vaillancourt	VP/CIO, Delaware Health Information Network
Michael Valle	Deputy Director/Chief Information Officer, HCAI - Office of Information Services (CA)
Patricia Vidal	Administrator, Agency for Healthcare Administration (FL)
Brenda Watson	Advanta Government Services, LLC
Christopher Whiteside	Whiteside Epidemiologist
Claudine Williams	Director, HDMI, State of Maryland - HSCRC
Herbert Wong	Director, Division of Statistical Research and Methods, Agency for Healthcare Research and Quality
Jodi Wooten	Georgia Office of Health Strategy and Coordination
Marian V. Wrobel	Principal Researcher, Mathematica
Narendar Yalamanchilli	CEO, Datagaps Inc.
Kenneth Yeates-Trotman	Director, Maryland Health Care Commission
Rong Yi	Principal, Data Science Modeling, Milliman
Ping Yu	Chief Growth Officer, Arbor Research
Amy Zhan	Manager, Manatt Health

Appendix - Detailed notes for some sessions

Notes from Session 2B - Identifying the Value of Your Data Resources

Dana Richardson, Jenn Toms
Through this interactive session we sought to help participants unleash the true potential of their data and unite data magicians to: (1) Identify the hidden value in data resources and uncover potential products and services that data can support, (2) Craft compelling narratives that showcase data's impact and how it translates to valuable solutions for potential data users, and (3) Collaborate with peers to gain fresh perspectives.

Participants broke into small groups to identify existing products and services for their data customers, selecting either one of the four topics along the data analytics continuum or one the topics of "Marketing Your Analytics" or "Promoting the Value of Your Analytics".

The data analytics continuum differentiates between the different types of analytics being conducted with data, and as data analytics increases in complexity, the further along the continuum an individual finds themselves.

At its foundational level, the data continuum begins with descriptive analytics, which can include extracts, reports, dashboards, scorecards, and benchmarking. By analyzing historical data, this phase of the continuum seeks to answer the question, "what has happened?"

The description analytics group emphasizes the foundational role of descriptive analytics in the broader analytics process. Key points included:

1. Question Definition: Begin with defining and refining the question multiple times to ensure clarity.
2. Data Selection: Identify data that can answer the refined question and transform this data into usable information.
3. Utility: Simple descriptive analytics and datasets help in making data more accessible and usable for customers.
4. Importance: Descriptive analytics are crucial as they provide a complete and accurate picture of past events, which is essential for understanding data customers and informing other types of analytics.
5. Experience: The group has practical experience in handling diverse data sets, such as state-level hospital data, APCD programs, ambulatory surgical treatment center data, and mental health data.
6. Continuous Improvement: Emphasized the need for new ideas, improved automation, streamlined processes, better data formats, and uniform definitions to enhance data linkage and overall effectiveness.

Further along the data continuum participants found explanatory analytics, which can include diagnostics, data mining, variances, and correlations. By analyzing data further to gain new insights, this phase of the continuum seeks to answer the question, "why has it happened?"

The explanatory analytics group highlighted the potential and importance of explanatory analytics in developing a new product to help employers and purchasers understand healthcare cost drivers. Key points included:

1. Product Focus: Developing an explanatory analytics tool for employers and purchasers to analyze healthcare cost drivers.
2. Market Opportunity: Identifying a need for a scalable, cost-effective tool accessible to more employers, especially smaller ones, with a neutral perspective.
3. Current Support: Acknowledged the existing support from various organizations and vendors but saw a niche for their product.
4. Challenges: Highlighted the need for resources, funding, policy changes, and education to develop the product.
5. Data and Trust: Emphasized identifying necessary data for employer-specific insights and building trust in the data.

In the next phase of the continuum, participants were asked to consider predictive analytics, which can include statistical analyses, risk scoring, and predictive modeling. This phase of the continuum seeks to answer the question, "what will happen?"

The predictive analytics group focused on the importance of predictive analytics across a variety of topics and highlighted several key points:

1. Target Audience and Problems: Identified various data consumers, including legislators, policymakers, providers, care managers, and the general public, and the specific problems predictive analytics could help solve for each group.
    o Consumers: Price transparency and cost predictions.
    o Legislators/Policymakers: Impact of laws and policies on state funds and resident outcomes.
    o Providers/Care Managers: Targeted prevention and population health management.
2. Applications: Discussed practical applications of predictive analytics, such as using data to manage health during the COVID-19 pandemic (vaccine distribution, herd immunity predictions, mass testing locations).
3. Quality and Accuracy: Emphasized the importance of accurate predictions and the need to communicate the limitations and volatility of predictions to end users. Highlighted the role of new technologies like machine learning and AI.
4. Tool Agnosticism: Advocated for being open to different tools and methodologies depending on the context and purpose.
5. Communication and Expectations: Stressed the need to clearly communicate that predictions are inherently uncertain and to set realistic expectations, especially in different scenarios (e.g., direct patient care vs. population health).
6. Building Trust and Buy-In: Note the importance of achieving accurate predictions and demonstrating success to build trust and gain buy-in from stakeholders, particularly legislators and policymakers.
7. Borrowing from Other Disciplines: Suggested learning from fields like sports betting, where predictive modeling is already mainstream.
8. Scenario-Based Modeling: Proposed allowing end users to input different parameters and explore various scenarios to better understand potential outcomes and engage collaboratively in the prediction process.

Prescriptive analytics seeks to provide the route which should be taken to reach a desired destination and asks "what should happen?" It is the final and most complex sort of analytics in

the data continuum and can include simulations, behavioral analyses, and decision and policy making support.

The prescriptive analytics group highlighted several key points:
1. Current Usage: Noted that no one was currently doing prescriptive analytics within the group, prompting a discussion on potential applications.
2. Scenario Analysis: Conducted "what if" analyses to evaluate decision-making retrospectively, considering how different actions might have led to better outcomes.
3. Target Audience: Identified vendor solutions and policy makers as primary customers for prescriptive analytics.
4. Methods and Challenges:
   o Discussed synthetic controls and the limitations of propensity score matching due to individual complexities.
   o Considered using individual risk scores as benchmarks.
5. Real-World Examples:
   o Kentucky Medicaid: Angela described using augmented synthetic control to assess the impact of moving to an MCO model on healthcare spending, finding that it did not save money.
   o Accorded Model: Frank explained a model that evaluates the cost implications of various interventions using condition-based propensity matching at the member level, taking into account demographic and condition interactions to provide specific cost benchmarks and assess intervention effectiveness.
6. Utility: Emphasized the potential for prescriptive analytics to inform strategies, price negotiations, and vendor evaluations by providing real-time, scalable insights into the effectiveness of different interventions and strategies.

The fifth group discussed "Marketing Your Analytics" and focused on strategies to promote the use of analytics tools and services, specifically a provider portal aimed at primary care clinician teams. Key points included:
1. Target Audience: Identified primary care clinician teams as the primary users, not patients.
2. Goals: Aimed to improve quality by allowing providers to privately compare their performance with peers within their organization using specific use cases to meet their goals.
3. Key Strategies:
   o Champion and Liaison: Emphasized the need for a dedicated champion to advocate for the tool and build relationships, along with a backup to ensure continuity.
   o Unique Use Cases: Suggested providing distinct use cases and challenges to demonstrate the portal's value to different agencies.
   o Privacy: Highlighted the importance of offering comparisons privately to maintain confidentiality and encourage usage.
4. Improvement Ideas:
   o Iterative Release: Proposed releasing updates in digestible increments, following the "bite, snack, meal" approach for easier adoption.

- o Feedback and Evaluation: Stressed the importance of gathering feedback and evaluating the tool's performance to ensure it meets providers' needs.
        - o Incentive Value Proposition: Recommended clearly defining the benefits for clients and the organization to encourage adoption.
        - o Benchmarking: Advised aligning quality metrics with existing standards to avoid redundancy and demonstrate clear benefits.
5. Summary: The marketing strategy involves communicating to primary care clinicians that the tool helps them improve quality privately, allows for peer comparison without revealing identities, and uses tailored use cases to meet organizational goals.

The sixth group discussed "Promoting the Value of Your Analytics" focused on strategies to gain recognition for analytics work, particularly in selling data to the commercial sector. Key points included:
1. Commercial Sector Focus: Discussed the potential for selling data to commercial customers, with Colorado serving as a successful model.
2. Case Studies:
        - o Pharma Group: Collaborated with Colorado to track a new drug's market uptake, medication adherence, side effects, and the duration of patient use.
        - o Latino Patient Access: Investigated potential delays in medication access for Latino individuals compared to white patients, aiming to use data to address and intervene in disparities.
3. Commercial Use of Data: Highlighted the potential for benevolent uses of commercial data, despite concerns about misuse. Emphasized the positive impacts and safeguards in place.
4. New Ideas and Product Improvements:
        - o Suggested exploring new ways to utilize and sell data.
        - o Emphasized the need for continuous innovation in creating valuable products.
5. Stakeholder Engagement: Identified stakeholder engagement as crucial for promoting the value of analytics. Noted that states vary in their capacity for engagement but stressed its importance in keeping stakeholders informed about the data's value.
6. Promotional Strategy: Recommended keeping the value of data and analytics consistently in front of stakeholders to foster engagement and recognition.

Notes from Session 3A - Data Sharing

Prompt #2 - Building Trust: Data sharing requires trust. What are some of the concerns you have heard regarding data sharing and how have you addressed them?

Concerns Regarding Data Sharing:

- **Data Quality:** Accuracy, completeness, timeliness (e.g., lab instruments)
- **Data Use and Misuse:**
  - Unintended exposure of misleading data
  - Sharing data with unauthorized entities
  - Data misuse or modification
- **Data Security:** Breaches, unauthorized access
- **Attribution and Documentation:** Lack of clear ownership or documentation standards

Strategies to Address Concerns:

- **Data Quality:**
  - Attestation of data by providers
  - Robust data documentation
  - Data Use Agreements (DUAs) with clear use cases and data dictionaries
- **Data Security:**
  - Secure storage and access controls
  - User training on data handling procedures
  - Regular audits of data use (random and mandatory)
  - Strong DUAs with enforcement mechanisms (e.g., removing access for non-compliance)
- **Transparency and Trust:**
  - Demonstrating value proposition to data providers
  - Clear communication about data protection measures
  - Certification of data quality and regulatory compliance

Challenges:

- **Scalability of Audits:** Balancing audit frequency with data volume
- **Enforcement:** Limited penalties for data misuse
- **Bad Actors:** Malicious actors who disregard agreements

Additional Considerations:

- Feedback loops for data providers
- Regular review and update of DUAs

Prompt #3 - Breaking Down Barriers: What are the biggest technical challenges around data sharing? What are reasonable solutions for these challenges?

Technical Challenges:

- **Data Standardization and Interoperability:**
    - Different data formats across systems and platforms
    - Lack of clear ownership and governance standards (e.g., OFAC compliance)
    - Challenges with physical data movement
- **Data Security and Access Control:**
    - Concerns about unauthorized access and data breaches
    - Implementing secure API gateways
- **Legal and Regulatory Considerations:**
    - Varying data privacy laws across states
    - Difficulty ensuring reciprocity and transparency in data sharing agreements
    - Complexities of Data Use Agreements (DUAs) and Data Sharing Agreements (DSAs)

Solutions:

- **Standardization and Interoperability:**
    - Adoption of cloud-based solutions (e.g., Azure, SEPP)
    - Development of common data dictionaries and formats
    - Technical assistance programs for data sharing
- **Security and Access Control:**
    - Investment in robust security protocols and encryption technologies
    - User training on data handling procedures
- **Legal and Regulatory Frameworks:**
    - Streamlining data privacy laws across governing bodies
    - Development of standardized DUA and DSA templates
    - Education and training on legal requirements

Additional Considerations:

- **User Education:** Importance of providing clear step-by-step guides and ongoing support
- **Mentorship Programs:** Fostering collaboration and knowledge sharing among data professionals

Prompt #4 - Key Communication: What is important to communicate in terms of data sharing efforts? How and to whom would you communicate these efforts on an ongoing basis? What methods would be most effective?

What to Communicate:

- **Data Updates:**
    - New data releases, schedules, important changes
    - Inclusion of forecasts and trend analysis
- **Transparency Initiatives:**

- ○ Data provenance (origin and history)
- ○ Data quality limitations
- ○ Public comment periods
- ○ Legal aspects (access, permissions, usage restrictions)
- **Value Proposition:**
  - ○ Explain the benefits and impact of data sharing initiatives
  - ○ Demonstrate value to data providers and users
- **Process and Procedures:**
  - ○ Data collection methods
  - ○ Redistribution processes
  - ○ Change management procedures (cost, requests)
  - ○ Issue resolution plan (urgency, communication scope)
- **Technical Information:**
  - ○ Data formats and standards
  - ○ Security measures
  - ○ Standardized interfaces and audit trails

Target Audiences and Communication Channels:

- **Data Users:**
  - ○ Provide regular updates through reports, dashboards, and presentations (quarterly or more frequent)
  - ○ Utilize platforms like SharePoint or direct communication channels
- **Data Providers/Owners:**
  - ○ Maintain open communication about data usage and decisions
  - ○ Demonstrate value proposition and ensure transparency
- **Funders and Policymakers:**
  - ○ Highlight the impact of data sharing on decision-making
  - ○ Communicate funding decisions and final outcomes

Ensuring Effective Communication:

- **Standardized Communication Practices:**
  - ○ Consistent templates, formats, and messaging across platforms
  - ○ Spokesperson(s) designated for clear and consistent communication
- **Education and Training:**
  - ○ Provide training materials on data management for stakeholders and public
  - ○ Educate users on data usage expectations and limitations
- **Feedback Mechanisms:**
  - ○ Implement two-way communication channels for feedback and issue resolution
  - ○ Convene advisory groups for ongoing discussions

Additional Considerations:

1. **Legal Requirements:** Clearly communicate legal aspects of data processing, storage, and access
2. **Patience:** Acknowledge the time it takes for information to be processed and shared
3. **Regular Updates:** Maintain updated information on platforms and documentation
4. **Accessibility:** Ensure communication materials are accessible to diverse audiences
5. **Telos of Need:** Clearly articulate the purpose and necessity of data sharing initiatives

Incorporating Additional Thoughts:

- **Merge with Medicaid:** Address the potential for merging data from the Veterans Administration with Medicaid programs for a more holistic view.
- **Will Back Rise:** Acknowledge potential cost concerns and emphasize long-term financial benefits.
- **Finance-based:** Frame communication around the financial value proposition of data sharing.

Prompt #5 - Building the Future: What do you hope data sharing looks like in the future? What is the potential for creating a sustainable and secure data sharing ecosystem?

Vision for the Future:

- **Standardization:**
    - Universal data formats, dictionaries, and data fields across organizations
    - Standardized data request processes and data quality metrics
    - National API Registry for seamless cross-state data sharing
    - Cost-effective and efficient data sharing infrastructure
- **Security and Privacy:**
    - Secure Data Services with robust access controls and encryption
    - Enhanced Data Quality Scores to ensure data accuracy and reliability
    - Minimizing data sharing needs through single source or peer-to-peer models
    - Expansion of Do Not Share (DNC) rules for increased data privacy control
- **Efficiency and Accessibility:**
    - Real-time data updates with streamlined data processing workflows
    - Centralized data repository with standardized interfaces for EHR, labs, PHR, and HIE integration
    - Automated data mapping, review, and extraction processes to reduce data sharing burden
    - Shared live collaboration features for data sharing, updates, and audit trails
- **Transparency and Education:**
    - Clear and comprehensive communication strategies for all stakeholders
    - Increased data literacy and education for providers, patients, and policymakers
    - Centralized location for evidence-based information on data sharing best practices

- **Sustainability:**
  - Uniform data policies and regulations across all states
  - Shared hardware and software demands to optimize resource allocation
  - Enforced data quality requirements for long-term data usability

Examples of Potential Benefits:

- Improved care coordination through seamless data exchange between providers
- Faster and more effective research through readily available, high-quality data sets
- Reduced administrative burden for providers by streamlining data sharing processes
- Enhanced patient engagement through secure, patient-controlled data access

Challenges and Considerations:

1. Balancing data standardization with the need for flexibility and innovation
2. Addressing privacy concerns and ensuring robust data security protocols
3. Achieving buy-in and collaboration from all stakeholders in the healthcare ecosystem
4. Addressing the cost implications of implementing a standardized data sharing infrastructure

Notes from Session 4A - Opportunities to Optimize APCD Operations

1. APCD Funding: Challenges and Opportunities

Current Funding Landscape:

- **Unreliable Grant Funding:** Limited grant opportunities with no guarantee of renewal.
- **Limited Revenue Streams:** Membership fees and data sales generate some income, but licensing is restricted.
- **Staffing Constraints:** Limited staff resources hinder grant writing and marketing efforts.

Challenges:

- **Loss of Major Funder:** Robert Wood Johnson Foundation funding has ended.
- **Data Submitter Concerns:** Limited data licensing options may discourage data submission.
- **Marketing and Communication Limitations:** Resource constraints hinder outreach to potential customers.
- **Grant Proposal Burden:** Securing grant funding requires significant staff time and effort.
- **Reaching Funding Capacity:** IAPD (Indirect Administrative Program Costs) funding has reached its maximum.
- **Underinvestment in Sustainability Efforts:** Limited resources dedicated to securing long-term funding.

Potential Solutions:

- **Subscription Services:** Introduce tiered subscription models for data access.
- **Membership Fees:** Increase membership fees to generate a more predictable revenue stream.
- **Targeted Marketing:** Focus marketing efforts on state agencies, hospitals, and universities.
- **Data Sales:** Expand data sales to other federal agencies.
- **Showcasing Value:** Host semi-annual data meetings to showcase the value proposition of APCD data.
- **Transparency and Planning:** Develop a transparent pricing portal with clear pricing breakdowns and advance planning documents.

Conclusion:

APCD requires a diversified funding strategy to ensure long-term sustainability. By implementing a combination of the proposed solutions, APCD can build a more predictable funding model and continue to provide valuable data services.

2. Analytics Pain Points: Challenges and Solutions

Challenges:

- **Data Quality Issues:**
    - Claim versioning and reconciliation lead to data inconsistencies.
    - Validation reports identify data quality problems, but require manual intervention.
    - Difficulty identifying and handling duplicate data entries.
    - Potential for concealed issues like "dead" or corrupt files during data submission.
    - Data copy issues can further complicate analysis.
- **Limited Analytical Capabilities:**
    - Current reliance on Tableau limits in-depth analysis.
    - Automation efforts face challenges.
- **Workflow Inefficiencies:**
    - Bringing analysts into the process later hinders early problem identification.
    - Lack of clear communication with data submitters can lead to data quality issues.

Solutions:

- **Improved Data Quality Management:**
    - Implement cloud-based versioning for better claim tracking.
    - Enhance validation processes for automated data cleaning and blocking of poor-quality submissions.
    - Develop strategies for flagging and handling duplicate data entries.
- **Advanced Analytics Tools:**
    - Explore additional tools like SAS and SQL Server for more sophisticated analysis.
- **Streamlined Workflows:**
    - Involve analysts earlier in the data processing stages to identify and address issues proactively.
    - Foster stronger relationships and communication channels with data submitters to improve data quality at the source.
    - Consider agile development methodologies for iterative improvement of data processing workflows.
- **Automation:**
    - Prioritize automation of data cleaning and validation tasks where possible.
    - Develop automated cross-tabulation for data control.
- **Stakeholder Feedback:**
    - Implement mechanisms to gather feedback from stakeholders on data quality and analytical output.
    - Utilize feedback to continuously improve processes and data usability.

3. APCD Data Management and Quality: A Summary

This document summarizes key points regarding data management, quality assurance, and stakeholder engagement within the APCD.

Ensuring Data Quality:

- **Data Submission Guidelines:** Clear and comprehensive Data Submission Guides are essential for ensuring data quality at the source. User-friendliness and ongoing updates are crucial.
- **Stakeholder Feedback:** Regular data releases with opportunities for stakeholder feedback provide valuable insights into data accuracy and usability.
- **Data Validation Checks:** Implement automated and manual data validation checks at intake and throughout the processing pipeline. These checks should focus on completeness, validity of values, thresholds, and logical relationships between fields.
- **Data Benchmarking:** Utilize data from state agencies (e.g., Division of Insurance, Medicaid, Public Health) to benchmark APCD data and identify potential inconsistencies.
- **Data Reviewer Expertise:** Assign data review tasks to individuals within the **Data Quality Unit** with a deep understanding of the data and the ability to identify anomalies.

Data Management Strategies:

- **Early Integration of State Files:** Embedding state data files early in the processing workflow can facilitate early identification of issues.
- **Flexible Data Processing:** Maintain flexibility to adapt to changes in data formats or submission procedures from payers.
- **Data Review Cadence:** Conduct data reviews at various time intervals (monthly, quarterly, annually) to monitor data quality over time.

Payer Management:

- **Consistent Communication:** Maintain consistent communication channels with data submitters (payers) through the **Payer Management Team**. This team addresses questions, clarifies expectations, and fosters collaboration.
- **Positive Reinforcement:** Focus on positive reinforcement strategies (e.g., recognition, collaboration) to incentivize high-quality data submissions.

Organizational Structure:

The APCD is composed of several key teams responsible for data management and quality assurance. These include:

- **Data Quality Unit:** Oversees data validation, benchmarking, and works with data reviewers to ensure data accuracy.
- **Payer Management Team:** Manages communication with data submitters (payers) and promotes high-quality data submissions.

- **Data Processing Team:** Prepares and processes incoming data files.
- **Analytics Team:** Utilizes the high-quality data to conduct analyses and generate reports.

Conclusion:

APCD prioritizes data quality through a multi-pronged approach, including clear guidelines, stakeholder feedback, data validation, benchmarking, and collaboration with data submitters. By continuously monitoring and improving data quality practices, APCD strives to ensure the accuracy and usefulness of its data products.

4A APCD Data Privacy, Security, and Governance Summary

AI Considerations:

- **Safety and Regulation:** Focus on maximizing the safety and ethical use of AI within the context of data privacy regulations.
- **Use Cases and Oversight:** Clearly define appropriate use cases for AI in data analysis, with human oversight to ensure accuracy and avoid bias.
- **Potential Risks:** Acknowledge potential security risks associated with AI and implement robust cybersecurity measures.

Data Governance and Security:

- **Standardized Data Request Process:** Develop a common data request application (ASPE) while acknowledging challenges in adoption due to state-specific regulations.
- **Data Sharing Protocols:** Establish clear policies and procedures for data sharing within the state, considering statistical de-identification techniques for broader data utilization.
- **Governance Skills and Consistency:** Identify the need for in-house expertise in data governance and strive for consistency across state data governance models.

Key Challenges:

- **Balancing Data Sharing and Privacy:** Finding the right balance between facilitating data sharing for research and analytics while protecting individual privacy.
- **Cybersecurity Threats:** Mitigating potential cybersecurity breaches through robust security protocols.
- **Harmonizing Data Governance:** Addressing inconsistencies in data governance practices across different states.

Additional Considerations:

- **Education and Training:** Importance of ongoing education for staff on data privacy and security best practices.
- **Evolving Regulations:** Staying informed about and adapting to changes in data privacy regulations like HIPAA and AI-specific policies.

This summary highlights the key areas of focus for APCD regarding data privacy, security, and governance. By addressing these challenges and considerations, APCD can ensure the responsible use of data for public health research and analytics while maintaining the privacy of individuals.

5. APCD Client Engagement and Management Challenges

Challenge: Clients Lack Data Expertise

- Users often lack the skills and knowledge to effectively utilize the data.
- Current documentation is perceived as inadequate for user needs.

Solutions:

- **Tiered Data Products:** Develop various data offerings tailored to different client skillsets (e.g., pre-processed datasets, basic vs. advanced analytics).
- **User Training Programs:** Implement training programs to equip users with the necessary skills for data analysis and interpretation.

Challenge: Communication and Expectation Management

- Ineffective communication leads to client frustration and unmet expectations.

Solutions:

- **Clear Communication Channels:** Establish clear and consistent communication channels with clients.
- **Data Use Agreements:** Utilize data use agreements (DUAs) to set clear expectations regarding data access, usage, and limitations.
- **Data Request Application:** Implement a data request application to streamline the request process and manage client expectations.
- **Transparent Data Queue:** Publish a data request queue to provide clients with visibility into processing timelines.

Additional Considerations:

- **Staff Training:** Ensure internal staff is well-trained on data acquisition processes and client communication best practices.
- **Data Quality Documentation:** Improve documentation to include clear explanations of data quality metrics and potential inconsistencies within the data.
- **Ongoing Engagement:** Develop strategies for ongoing client engagement to ensure user satisfaction and continued data utilization.

By addressing these challenges and implementing the proposed solutions, APCD can build stronger client relationships and empower users to leverage the full potential of APCD data products.

6. APCDs for Policy and Data Efficiency

Challenge: Policymaker Disconnect

- Policymakers may lack a clear understanding of APCD data capabilities and limitations.
- Difficulty in amending existing data governance structures due to their rigidity.

Solutions:

- **Stakeholder Engagement:** Establish a committee with representatives from policymakers, data users, and APCD leadership to facilitate communication and inform policy development.
- **Data Governance Review:** Conduct periodic reviews of data governance policies to ensure they remain aligned with evolving data capabilities and user needs.
- **Policy Flexibility:** Encourage policymakers to enact flexible data governance frameworks that enable APCD to adapt to emerging trends.

Additional Considerations:

- **Data Literacy Initiatives:** Invest in data literacy initiatives to educate policymakers on the potential and limitations of APCD data.
- **Collaboration:** Foster a collaborative environment between policymakers and APCD to ensure data-driven policy making.

This summary emphasizes the importance of bridging the gap between policymakers and APCD governance. By implementing the proposed solutions and fostering ongoing collaboration, APCD can ensure its data resources are effectively utilized to inform sound policy decisions.

Notes from Session 5A - Governance is Critical to a Successful Program

Session 5A Notes
Governance is Critical to a Successful Program
Facilitators - Angela Taylor and Craig Schneider

Discussion Topics
- Data Governance Experience
- Creating and Maintaining Data Governance
- Evolution and Adaptation
- Balancing inclusivity vs. efficiency

I. Building and Maintaining Data Governance Programs
- Start-up Phase
  - What steps should HDOs take to establish a data governance program?
- Operational Phase
  - What ongoing tasks are there for HDOs regarding data governance?
- Recent Changes
  - How has data governance evolved in recent years? (e.g., increased security protocols, legal scrutiny)

II. Curating and Maintaining Data
- Data Sources
  - Where can you find relevant policies, forms, and best practices? (e.g., looking at other states)
- Data Release
  - What data can be released, to whom, and how? (e.g., legal considerations, data use agreements)
- Post-Release
  - How do you follow up on data usage? (e.g., data destination, audits, publications)
- Recent Changes
  - How have data release policies changed due to legal and technological advancements? (e.g., Roe v. Wade decision)

III. Balancing Inclusivity and Efficiency
- Approaches to Data Access
  - Examples of how states balance data access with efficiency (e.g., standardized DUAs, tiered access)
- Prioritization
  - How to prioritize data requests while ensuring inclusivity (e.g., tiered triage processes)
- Cost Structure
  - How can data access costs be structured to be inclusive? (e.g., dependent on requester type)

IV. Evolution and Adaptation

- Adapting to Change
    - How should data governance adapt to new trends, technologies, and regulations?
- Examples of Adaptation
    - Specific examples of how data governance practices have changed in response to external factors.
- Building Flexibility
    - How can data governance be structured to be flexible and adaptable in the long term?

V. Data Governance Experiences
- Successful Approaches
    - What practices have been effective in implementing data governance?
- Problematic Approaches
    - What challenges have arisen in implementing data governance? (e.g., bottlenecks, data quality issues)
- Lessons Learned and Best Practices
    - Key takeaways and recommendations for effective data governance.

VI. Additional Discussion Points
- Bottlenecks in the data release process (e.g., legal approvals)
- Ensuring timely data access
- Maintaining data quality standards
- Sharing best practices for data governance documentation