NAHDO In-Person Workshop Report and Session Summary

May 10-11, 2022
Seattle, Washington

Common Themes

- Data Quality
    - Databases continue to have gaps and quality issues that need to be addressed
    - Missing fields and data lags continue to limit the potential of the data
    - As new questions arise, the need for addressing data quality is becoming more important
    - Data programs are facing pushback on collecting person-identifying information, which requires creative solutions
    - There is a wealth of potential in linking data from multiple sources
    - With the significant interest and investment in interoperability and data sharing, there are many needs for standardization of data, policies and processes
    - We have a long way to go to get data consistency across states for claims data
- Changing Payment Models
    - It is inevitable that payment models will change
    - We need to be capable of adapting data collection and analysis to address those changes
- Data Governance
    - HIPAA is outdated and needs updating
- Medicaid Unwinding
    - The rise in the number of Medicaid enrollees due to the emergency rule has shifted trends and the unwinding of that bump in enrollment will likewise cause issues that will need to be addressed
- Equity Issues
    - There is an increased interest among stakeholders and policy makers in examining and understanding equity issues
- Responsiveness
    - As the number of use cases for health data are rapidly expanding, it is important to meet the data and information needs of various stakeholders, especially policy makers
    - Continued support for data collection and use will rely on our ability to show value in what we do
- Funding
    - Progress in many dimensions is resource limited, so some form of focused funding will be required to make real progress
- The Power of States
    - The collective experience of state health data organizations is vast
    - NAHDO members should not hesitate to ask each other questions and share experiences

Deep Dive on Data Linkage

Facilitators - Andrea Clark, BerryDunn, and Kenley Money, ACHI

Introduction

Health data organizations regularly link data. There are many data sources and methods for linking data to improve completeness and perform analysis. A common example is linking a person's records over time and across services. However, not all data linkage addresses person identity or uses direct identifiers. Many researchers use geographic aggregates to better understand aspects health care using another data source (e.g., air quality).

Session Objectives

1. Understand the linkage challenges and opportunities with administrative and non-administrative data.
2. Share best practices and challenges related to administrative and non-administrative data linkage activities.
3. Participants should leave with a mid-to-high level understanding of administrative and non-administrative data solutions other participants are using and be able to apply these solutions or components to their own data and analytic needs

---

Session Takeaways

Key thoughts

Complexities and Challenges

- Legal challenges
    - Data governance
    - HIPAA is outdated - we need new tools to help alleviate HIPAA concerns
    - Re-identification risk
- Data Issues
    - No consistent definitions for race and ethnicity codes
    - Provider side of service
    - Need for specific provider attribution methodology
    - Lack of master provider file
    - How to distinguish between divisions and departments in hospitals
    - How to distinguish residents from hospitalists
    - Missing data elements/data gaps (need master patient id?)
    - SSN not being required anymore
    - Clear text, non-standardized values for some fields
    - Timeliness – when data is collected vs when it is used
    - Lack of uniformity for coding, hospitals
    - Data limitations based on varying acceptable margins of error
    - Differences in EHR data vs claims
- Issues with Person Identifiers
    - How to truly identifying individuals. Finding the correct data for matching.

    - Linking claims to inpatient discharge and emergency department data
    - Tracking across plans and services
    - Ethics behind datasets being linked
    - Multiple people associated with a birth (need family ids?)
    - How to create the final or ‘golden’ record? When to use probabilistic matching vs deterministic

Best Practices and Solutions

- Develop and share standards across carriers and states for hashing and similar algorithms
- Providing data aggregators with endorsements
- Utilizing HIPAA compliant tools – FHIR, HL7
- Best tools for connecting services and individuals
    - Episode Groupers
    - Categorization tools – CCS, NCPDP
    - CDL
    - Multi-state modules
    - Inclusion of self-insured data

Unresolved Issues that Require Further Methodology Discussions

- Master Provider File solutions
- Tracking individuals across plans and services
- How to handle missing data such as SSN
- How does HIPAA need to be refreshed

COVID-19 Data

Facilitators - Betty Harney, CHIA, and Kenneth Geter, NM Department of Health

Introduction
Two years into the COVID-19 pandemic, many data issues / challenges have already been encountered and addressed. We need to discuss these to identify what we can learn from and what we need to do going forward based on those lessons.

Session Objectives

1. Attendees will share knowledge and experiences and challenges we've faced and the work we've already successfully completed for the public good
2. Attendees will identify remaining challenges and we should collaborate on possible solutions.

Session Takeaways

Key thoughts

Intake

- Missing fields - COVID data for testing and vaccines may come in one file from the public health agency, but many fields are not complete. This was largely due to the short turnaround (every week) of the data submission and nobody was taking the time to do audits or edits. While some of the gaps will be filled later as part of a fill and replace submission, it is not always the case.
- Lags in claims data - Information on testing and vaccines is near real time, but the rest of the data is still on a significant delay, which can be up to six months.
- Different source data permissions (lab and testing data did not always align with the ability to look at claims data)
- What happens to the zero-claims data? Are they showing up in the data or are they disappearing?
- Medicaid Unwinding - What is that going to do to the volume of claims data? Right now, the number of covered lives is perhaps at an all time high, but that will slowly decrease over time.
- There was some consolidation of data which can be very disruptive.

Data Quality and Release

- People changed payer types - to Medicaid or uninsured - due to COVID and we are not sure where that data will be coming from.
- There are new streams of data coming in and not everyone will have access to all of those streams.
- Where will vaccines show up? Are they going to be pharmacy claims, medical claims or something else?

Health Equities

- Different schemas for identifying disparities and gaps - ADI, Social viability index, food security index. Each has its own particularities.
- There are challenges to quality that are being exacerbated by COVID. Quality metrics are starting to deteriorate. This might drive greater gaps in quality and outcomes.
- There are also challenges in self-identification of the population. Different states have different granularity in the ways identities are defined.
- There are risks to patients related to privacy and risks to research due to not having sufficient data due to suppression.

Data Needs of Policymakers

Facilitators - Leanne Candura, HSRI and Kathy Hines, CHIA

Introduction

In the realm of health care, there are a wide variety of people who form ideas and make plans that impact the system. This includes individuals in both government and healthcare industry situations. At all levels of policymaking, having access to complete, correct, and timely information is highly beneficial in supporting overall policy goals. There is so much data that is already being collected; we just need to figure out how to put it to work.

In this session, we will focus on the "art of the possible" - how can we move from a world where policymakers are sometimes suspicious or even hostile toward data to one where we jointly celebrate the great things that can be done?

Session Objectives

1. Attendees will discuss the disconnect between what policymakers think they need to know and what data people think the policymakers ought to know.
2. Attendees will formulate ideas and action plans to provide policymakers with better information and data.
3. Attendees will leave with at least one action item that will better support policymakers.

Session Takeaways

Key thoughts

1. Communication is key - know what the question is, who is asking the question, have the conversation about whether it can be answered/what is available, and offer an alternative as to what can be provided
2. You need someone skilled in government relations that knows the policy maker landscape. They should be experts on the process. They can help you avoid landmines and understand the story behind the question.
3. You can build credibility in your relationships so that they will accept the fact that you can't always answer every question.
4. Your members or supporters pay for more than just developing data. They want you to drive change
5. Be proactive - developing use case libraries, dashboard on common themes or other tools will help policy makers understand how the data can be used more appropriately. Anticipating what the priorities and focus in the near future (such as for the upcoming session) can guide development of data resources.
6. Pivot if data isn't available to get time/funding/support to collect the data

Working with State Legislators

Often, they know what data they need and the question is whether they are willing to match that up with the data that you have. They need to understand that the data that they really need is often the complement of the data that we have. Questions may have to be reframed to be answerable, yet still useful. Make sure you are transparent about what you have and what you can do. You can turn this into an opportunity to advocate for additional data access or resources.

There are tools you can develop that are helpful. In most states, there are common themes that tend to be persistent. In those cases, data dashboards and standardized reporting may not answer all of the questions, but they can help frame the questions so that they are more better informed and more concretely framed. Developing the ability to generate comparative analysis by legislative district can be a very helpful tool. These can be in the form of template reports that give qualified results about what the data is and what it isn't.

Many requests are intertwined with the needs and support of advocacy groups. If there are opposing groups on an issue make sure you are transparent with the information you provide so that there is no perception that you are favoring one side on the issue

Working with the Executive Branch

They need quick, current data tailored to particular decision points. They have a particular problem at the moment they are trying to decide. Live dashboards can be helpful, but have significant limitations. There is constantly a challenge because of the difficulty of tailoring the output to the question. Often the questions are not clearly defined and we have to reach beyond what they are asking to figure out what analytics they need. Try to narrow down the theme to a more guided data request. This will cut down on work but more importantly help ensure their needs are met with data you can stand behind - aligning the need with what is available.

Soft skills are really important in our organizations - how to shape questions and make them answerable. Understanding the limitations of the data but still making it useful and valuable. Staying on solid ice.

Interoperability

Facilitators - Josie Bechara, NORC and Lisa Bari, Civitas Networks for Health

Introduction

In the context of health data, interoperability refers to the ability of data programs and systems to share or exchange data with each other or with individuals, such as providers or patients. Successful interoperability requires conversations about data architecture, interfaces, and standards that allow sharing or exchange in an appropriate, secure manner.

Session Objectives

1. Attendees will have a broader understanding of interoperability and how it can impact their data program
2. Attendees will identify important use cases for the sharing or exchange of health data
3. Attendees will develop a plan of action for improving interoperability

---

Session Takeaways

Key thoughts

- Interoperability is the ability to connect data from different sources based on common definitions and standards using curated, secure technology
- Incentives and value propositions across organizations
  - Quality of care and quality of data
- Open standards will support the idea that no one data vendor/holder has too much control in space - help democratize the landscape.
- Use cases will highlight the importance of value proposition to all data sharing organizations (what's in it for them)
- Make sure that patients have own access to their own data
- One barrier to interoperability - Fear of how data will be used/processed
- Context is everything - workflows vs. actual data needs
- Culture is also important
- Integrating data from multiple sources requires interoperability in one form or another
- Legislative limits on governance can present challenges to interoperability that may be subject to influence of changes in administration
- The technical challenges are at least knowable - The bigger challenges are related to governance and legislative limits
- Federal agencies - ONC, CDC, CMS, HRSA - have a role in providing technical guidance as well as opening up possible financial resources. Investment should be at the state level, which will provide more benefit when there is consistency at various state and federal jurisdictions

Key Recommendations

- Ensure that there is a ton of transparency about how the data will be used and shared
- Standards and definitions are essential for interoperability
  - Data Elements

    - Data Exchange
    - Data Governance - Having an umbrella governance structure would be very helpful and might require creating a unified data space where all health data “live together.”
    - Workflows
- Important to have strong relationships that continually show value and show agencies how to be interoperable
- Data organizations (APCD, HIE, QIO, etc.) need to present a unified face among related groups/partners and create “health data spaces”

Multi-State and Regional Data Projects

Facilitators - Scott Leitz, NORC, and Kyle Russell, VHI

Introduction
The majority of health data collection systems are administered at a state level. Several types of analyses have the potential to be even more impactful when state boundaries are not taken into account. For this to become a reality, a wide variety of pitfalls and potential barriers must be accounted for. This session aims to explore some of the obvious, and not so obvious, key factors in driving the success of future initiatives.

Session Objectives
1. Determine the barriers most likely to detract from multi-state data collaborations and potential solutions
2. Formulate key criteria for multi-state collaborations to be successful
3. Identify the types of analyses most likely to benefit for these collaborations

---

Session Takeaways

Key thoughts
Multi-state Collaboration
- Challenges that cause implementation to be difficult or fail
  - Legislation that governs the process may not be consistent across states
  - Providers might not want shared across state and resist the effort
  - Health plans could be similar to providers in concerns
  - Budget or funding may not be equal across states and could be cut on short notice
  - There can be issues related to competition or rivalry
  - Mutual trust - decision making and process needs to be agreed upon, including any restrictions on governance, privacy and security
  - Changing priorities, including changes in leadership or administration
  - There may be a variety of data issues, including technical issues and linkage issues, incompatible forms or data elements
- Features that will make it easier
  - Common stakeholders or experiences where states already invested together, easier to coordinate
  - There can be financial incentive or mutual ROI to participate
  - Legislation might bleed over and be similar state-to-state
  - States may experience FOMO and not want to not be left out
  - Commonality is appealing and beneficial (early investment in the states)
  - A common vendor across the participating entities will make it easier to pull the data together; handle data commonalities
- Why some states may not get on board with a common data layout

    - There is always a cost associated with the rework required to change to a new layout
    - State-specific statutory requirements and/or legislative rules require data elements that exceed the common layout (state needs fields not in the layout)
    - Political forces/populations and subsequent needs represented within each state differ; Enough political support may not be available
    - A year may not be enough time to develop/come to an agreement on a common layout
    - Each state has different tolerance levels pertaining to privacy and security
    - Not enough funding resources for the planning/research period required for developing a common layout
- Ideal topics for multistate collaboration
    - Opioid Epidemic
    - Maternal Mortality
    - Exposing geographic variation in more complete way
    - Independently assessing what is like to be a Medicare/Medicaid beneficiary in different states
    - Fraud, waste, and abuse across states
    - Cybersecurity
    - Shared understanding across states in building and running APCD
    - Hospital payment comparison
    - Cross state lines and cross state care
    - Substance use and adverse effects
    - Follow a person across the states
    - NCSL - identify leaders to get multi-state leadership across SUD and other issues
    - Make it easier to get data requests across multiple states
- Biggest Barriers to multistate collaboration
    - No comprehensive federal law
    - Lack of consistent funding
    - Alignment on standards
    - Lack of common goals
    - Agreement on common approach
    - Politics
    - Priorities are local and may not translate
- “Magic Bullets” that would make multi-state projects happen
    - People start caring enough that they are willing to work together at federal and state level to get it done
    - Grants that specifically target making multi state work possible
    - Common ground on resources, laws, procedures for data sharing

---

Non-Claim Payments

Facilitators - Alice Aguirre, CIVHC, Janice Bourgault, Onpoint Health Data, Amy Kinner, Onpoint Health Data and Charles Hawley, NAHDO

Introduction

Many states collect claims data to support the triple aim of improved population health and quality of care and reduced costs. However, claims data do not always capture non-fee-for-service payments made to a provider. The RAND Corporation and Milbank Memorial Fund have published guidance on measuring the total cost of primary care and the need for non-claim payments data to do so accurately. Several states have even begun collecting and using non-claim payments data as part of their APCD programs.

Session Objectives

1. Participants will develop a better understanding of the need for and use of non-claim payments data.
2. Understand existing efforts to collect data and produce information on medical costs and drug rebates.
3. Understand initiatives exploring commonality in data collection and the role of standardization.

Session Takeaways

Key thoughts

- Total Cost of Care
    - Non-claim payments are defined as anything that is not a fee-for-service payment
- Data Collection Issues
    - Long lags - Many of these payments are made months after the end of the year, so you may need to get preliminary and final data.
    - Encounter data to accompany non-claim payments is desired but is problematic
    - Submission errors - Plans can find errors after the attestation date
    - Definitions are not uniform - How do we ensure all payers interpret them the same?
    - Trade secrets - Granular data may be protected
    - No visibility into what the payers are actually submitting - difficult to validate
    - Claims data and APM data may be coming from two different sections of the payer’s system that may not communicate
    - Claims data are often based on the residence of the patient, but APM data is based on the situs of the provider or possibly where the plan is issued
    - APM payments very likely include self-insured payments along with fully insured
- Drug Rebates
    - Understanding patient responsibility and trends over time
    - Two states currently doing this - CO (getting data by therapeutic class) & MA
    - How are coupons to be treated?

    - Part D has limitations on the granularity of the data
    - Detailed rebate information is needed to do trend analysis because it is such an important part of the equation
    - The ideal would be to get more detailed rebate information that can be tied to particular drugs or even members
- Comparing Across States
    - Data is not consistent state-by-state
    - The CDL doesn’t really look at these issues yet and will need to include something on rebates and alternative payments to help fill the gaps - necessary for moving the field forward
    - Nobody is going to lose their competitive advantage over the collection of this data - this would allow providers to have benchmarks

---

Analytics to Inform Public Health

Facilitators - Angela Taylor, Commonwealth of Kentucky and Nicole Sidrak, VHI

Introduction

Public health is an important and very broad segment of the health community. Public health workers promote and protect the health of people and the communities where they live, learn, work and play. The success of public health relies on accurate and timely data and careful analytics. Health data organizations can play an important role in ensuring that public health has access to and is capable of using data from a variety of existing sources.

Session Objectives

1. Each attendee will have a new awareness of how data and analytics can be used to inform public health activities
2. Attendees will have a broader understanding of data sources and analytics that are currently available to inform public health
3. Attendees will identify barriers that prevent public health from using data and analytics and propose solutions to overcome those barriers

---

Session Takeaways

Questions:

1. What data is available?
2. What barriers prevent public health workers from using data in this area?
3. What can be done to overcome those barriers?

Group 1 - HIV

1. APCD (pre-exposure prescriptions); HIV registries (infected); youth behavior risk survey; BRFSS
2. Data that goes into registries comes out aggregated; data security; RX for treatment and for prevention are the same
3. Rhode Island; 2 vs 3 RXs; look for areas not utilizing RX

Group 2 - Cancer

1. Registries, Hospital discharge, Clinical records, Vital records, APCDs used for Strategies, treatment opportunities, analyzing disparities, screenings, research purposes
2. Data available casts such a wide net that the people using the data might not actually know what's available; prioritizing what to invest your time/resources in - beneficial to have a focused objective/start small and then build, gap between people doing clinical interventions and those performing the analytics - both groups need education; too many data sources to clean/verify quality
3. Tapping into expertise; prioritizing; having a focused objective; education of public health population pertaining to data sources available; transparency regarding successful use cases

Group 3 - Child Welfare

1. Child welfare data system as required by federal law; Medicaid/APCD data (foster kids are usually Medicaid eligible); education data/school outcome data; vital records;

juvenile justice system data; discharge data; BRFSS; hotline data; public health registries; environmental risk data; SDOH; food deserts

2. Siloing/extreme sensitivity of data; PR/lack of awareness and/or open conversations; people on the ground with the families (social workers, etc.) are not data-driven people and may not have the time to dedicate to getting the data where it needs to be (resource issue)
3. Take baby steps; build aggregate dashboards to publicize quantity of kids in foster care, common outcomes, etc. to bolster awareness; trying to make data capture as easy as possible (online/offline apps); forming multi-disciplinary teams

Group 4 - Infectious Diseases

1. Data collection for required surveillance/case reporting systems; claims data and discharge data to determine burden of certain diseases; RX data; social media data; syndromic surveillance
2. Data collection is provider-driven and time consuming; data access; expense associated with text mining; HL7 technology not being prioritized
3. Increase funding for data systems; enforcing standards; enforcing requirements properly and making providers aware of requirements

Group 5 - Maternal and Infant Mortality

1. Birth/death records; APCD claims; vital records; census info including SPI; hospital ER/discharge data; BRFSS data; social determinants; HMIS
2. Sensitivity/privacy; availability; linkage; time involved to track and link data; limited patient history; variations in how race/ethnicity data is noted across sources; financial constraints
3. Stratify data; create data models when unable to link 100%; develop method to standardize race/ethnicity; track/assess/educate/assist in pre-natal and post-natal

Group 6 - Substance Use Disorder

1. APCD; EHR/HIE; death records; corrections data
2. Linkage; federal regulations around receiving/reporting SUD data; regulatory ambiguity; data is biased due to different interpretations of federal laws and/or disproportionate effects on uninsured vs. insured populations; stigma of self-reported data; lack of APM data
3. SAHMSA to create clear guidelines; improve understanding/collection of APM data; use of RX data; incentivize better reporting of procedure codes/modifiers for screenings

---

Analytics to Inform the General Public

Facilitators - Cecilia Ganduglia Cazaban, UT Health, and Marcos Dachary, Milliman MedInsight

Introduction
States often have public reporting requirements to produce information for the general public that depends on analytic tools. The general public is often considered to include lay people as well as policymakers, other state agencies, researchers, and health community stakeholders. Health data organizations play an important role in ensuring information produced from a variety of existing data sources are accessible, interpretable, and useful to the general public.

Session Objectives

1. Each attendee will have a new awareness of how data and analytics can be used to inform the general public
2. Attendees will have a broader understanding of data sources and analytics that are currently available to inform the general public
3. Attendees will identify barriers that prevent the general public from using data and analytics to propose solutions

---

Session Takeaways

General Public Needing Services

- Proposed Idea: 411 for Health Data - bringing together resources and data to empower patient advocacy
- People need to know what data exists and where to access it
    - Patient rights
    - Costs of procedures
    - Networks and coverage
    - Procedure quality and volume at different providers
    - Referral resources and connections to insurance
    - Readmission rates
- Pulls together information from datasets about specific facilities and providers
- Set of questions that patients or their advocates should be asking their provider

General Public with Specific Conditions

- Proposed Idea: Design an app that meets the needs of specific conditions
- Allows access to many diseases, such as substance abuse or opioid risk
    - Risk factors specific to your situation or geography (conditional logic based on input and answers to questions)
    - Access to treatment near you
- Can also be useful for family or friends that have an interest in understanding the patient’s condition
- Chat function - get immediate feedback

- Must be scalable - start with one condition and built out to many others

Information in Public Health

- Proposed Idea: An app for children
- Children need access to age-appropriate information about themselves and their health
    - Blood type
    - Vaccination records
    - Dental exam schedule
- Teach children about health care - can be gamified
- Allows public health to push information, such as warning about dangerous behaviors, like vaping
- Can be a catalyst that will create informed patients in the future
- Must have age controls so that only appropriate information is accessed and shared

Employer Community

- Proposed Idea: A report suite or tool kit for employees of smaller employers
- Information from your health plan that is important but doesn't exist: Who did the most procedures and who had the best outcomes?
- Electronic and paper reporting are both needed
- Targeted quality metrics that would play a big role in consumer decision making

---

Geographic and Location Data

Facilitators - Victoria Udalova, Census Bureau and Caig Schneider, HMA

Introduction

Geographic and location (GL) data can be used along with individually identified data and other sources to provide valuable insights in many dimensions. The geographic data may be fundamentally interesting, such as for exploring differences in outcomes for urban or rural populations, or the geographic data may allow linking to data sources that represent proxies for unknowable information at the individual level, such as poverty status and social challenges.

Session Objectives

1. Participants in this session will gain a better understanding of both the applications and uses for GL data and the value and limitations of potential sources.
2. Participants will develop strategies for incorporating GL data into their analytics in order to answer new classes of questions that rely on geography or location.

Session Takeaways

Applications & Uses

Challenges

- Geographic areas and definitions don't always easily line up across data sources
- There is value in mapping your datasets to other sources, such as census, school districts, legislative districts, etc.
- Mapping to tribal geographies, such as reservations can be difficult
- Location of subscriber may not be location of everyone on plan (college kids, for example)
- Sometimes you need to combine multiple years of data to get to small geographies
- Timeliness can be an issue
- What would it take to get better participation?

Use Cases

- Geo-location of COVID testing and vaccine sites
- Air ambulance use by geography
- You can connect ACS data to claims data
- Tech companies are collecting a lot of geographic data for application/marketing that could also be used in connection with health data
    - Food deserts & health care sites
    - Apple, Google, Uber, etc. know a lot about where people live, work and how they get transported

Potential Sources

Census Bureau Data

- The ACS contains 5 year data which has a publicly available component.
- More granularity is available through Census Data Centers, but access is restricted.

- It would be good to provide information about how this access happens.
- Colorado - their vendor gives them access to census tract data, linking to the ACS for housing and income

Other

- New Mexico - air quality indicators
- Be cautious about what is releasale
- Determination of urban vs rural though definitions not always consistent or easy to address
- Consideration of other processes to impute proxies for race and ethnicity using address, surnames or other issues

Best Practices

- The level of data you need depends on what you are trying to do - 3 digit/5 digit ZIP code, census tract
- You should consider geographic conversions and aggregation over time
- There is an important discussion that needs to happen regarding the value of increased granularity vs the need to protect privacy through suppression
- There should always be a step for provider engagement as part of the validation process
- Taking geography with grain of salt because there are known issues, and you should become familiar with them

Indicators and Indices

- Benchmarks tied to census data would be helpful
- Some level of grouping or classification of counties based on similar characteristics could be useful

---

Understanding Demographic Data (REaL, SOGI, etc.)

Facilitators - Chris Krawczyk, HCAI, and Tanya Bernstein, Freedman HealthCare

Introduction

The COVID-19 pandemic revealed gaps in our public health and health data infrastructure, the ways health inequities are perpetuated, and the need for a modernized public health data infrastructure. Demographic and SDoH data are critical to understanding disparities in health outcomes and making effective policies to address inequities. Modernizing the often disjointed health data infrastructure requires the support of the national, state, and local levels of government, private, and civil society stakeholders in public health, health care delivery, and health data systems.

Session Objectives

1. Participants will have a better understanding of the utility, availability, and quality of demographic data in the health data ecosystem.
2. Understanding tactical approaches states take to address the need for demographic data given the limitations of standards.

Session Takeaways

Key thoughts

Race, Ethnicity, and Language

- Sources: Enrollment data and eligibility data, HIE, Hospital Discharge Data
- Key Considerations
  - Cell suppression sometimes requires REL to be removed
  - Free text fields will require cleaning and validation
  - Sensitivity - people may not be willing to disclose this information on a form
  - A general reticence to engage on this topic
- Data quality issues
  - If the field is not required, it may suffer from completion issues
  - Outdated methods for collecting the data may lead to confusion
  - Language presents a lot of issues because people have had bad experiences with the system
- Mitigation and Resolution
  - Do you use the most recent indication or the reporting that provided the most granularity?
  - New methods similar to G&L data - imputation of language
  - Private data sources may provide assistance
  - Engagement with populations about how the data are collected and used

SOGI - Sexual Orientation and Gender Identity

- Sources:

  - Sex at birth and gender identity - birth certificate data, driver's license, insurance enrollment, HR data
  - Sexual Orientation - not systematically collected across the board
- Key considerations
  - Why do we need to collect this data? Should we be collecting SO data? Very helpful in some public health and behavioral health settings.
  - For screening purposes, you need gender to be effective at sending out reminders.
  - For coverage, it may also be relevant
- Data quality issues
  - This is a moving target. It is important to attach a timestamp and understand that it could change at any time.
  - There is no standardized way to collect the data into grouping. How do we make sense of the way it is collected? How do you distinguish between unknown, other and declined?
  - Sex at birth - Male, Female, Intersex, and Unknown
  - Gender Identity - Male, Female, Nonbinary, and Unknown
  - Transgender can be inferred from the combination of sex at birth and gender identity
- Mitigation and Resolution
  - How do we make this more appropriate? There will be lots of variability in how communities want the data to be collected - word choice, etc.
  - We need to figure out the balance between security and privacy making it more comfortable to be collected but more difficult to be used.

Social and Community Determinants of Health

- Sources:
  - ADI - Area Deprivation Index
  - SVI - Social Vulnerability Index
  - Social Determinants of Health Database (NORC created for AHRQ)
  - Census Bureau - Community Resilience Estimates (Tract level estimates of the number of people that have a given number of risk factors)
  - Area Health Resource File
- Key Considerations: These are looking at community-level data, instead of individual data, including Housing, Transportation, Food, Air & Water Quality, Crime
- Data Quality
  - Use caution to protect identity
  - Balance social needs with privacy laws (outdated HIPAA)
  - Need standardization of community assignment
  - Limits on availability of data
  - Tracking people that relocate or have no location on record
  - Accurate application of indices
  - Lost or unidentified population (uninsured, undocumented, mental health, etc.)

- Mitigation and Resolution
  - Need adequate geocoding
  - Master provider index and master patient identity index
  - Updated policies on privacy protections
  - Provider data to help track the lost population

<u>Health Care Access</u>

- Sources:
  - Eligibility and Enrollment (Claims data)
  - Network Adequacy and Provider Access (Claims data)
  - Healthcare Literacy
  - Medical Cost Burden (Claims data)
  - Access to Transportation
  - Telehealth Utilization (Claims data, broadband access)
- Key Considerations
  - A lot of these data are based on geography
  - Make sure you have the specificity that you need (noting that ZIP codes are awful to work with)
  - Some of the mapping aspects of transportation exist
  - Data on volume analysis, accepting new patients, and wait time to get in to see a provider might exist in some form in the states
- Data Quality
  - Don't address well the uninsured
- Mitigation and Resolution
  - To really leverage the claims data (eligibility and enrollment)
  - Wait for the Gravity Project to bake some of these elements into data
  - Look for geographic aggregate data

<u>Socioeconomic Variables</u>

- Sources:
  - IRS
  - ACS
  - School district (free and reduced lunch)
  - Lexis-Nexis
  - Credit Scores
  - Social Media mining
  - ACGs
- Key Considerations
  - High deductible health plans may impact the ability to access care
  - Sensors
- Data Issues
  - Potential for abuse and running into privacy laws
  - Diversity is a challenge
  - Clinical vs. Public health measures and how they might work