Data Visualization

November 11, 2016
12:00 p.m. – 1:15 p.m. ET
Meeting Information

▪ Conference Line: 1-866-269-6685
▪ Conference Code: 6763836672#
▪ Reminders:
  ▪ Please **hard-mute your computer speakers** and the speakers in the web conference
  ▪ Please **mute your phone line** when you are not speaking to minimize background noise
▪ Technical difficulties? Email us at chpinfo@academyhealth.org
Chat Feature

- To share your comments using the chat feature:
  - Click in the chat box on the left side of your screen
  - Type into the dialog box and click the send button

- To signal to presenters you have a question / comment:
  - Click on the drop down menu near the person icon and choose *raise your hand*
DASH and CHP are All In!

Community Health Peer Learning Program

- NPO: AcademyHealth, Washington D.C.
- Funded by ONC
- 10 Participant and 5 Subject Matter Expertise communities

Data Across Sectors for Health (DASH)

- NPO: Illinois Public Health Institute in partnership with the Michigan Public Health Institute
- Funded by RWJF
- 10 grantee communities
All In: Data for Community Health

1. Support a movement acknowledging the social determinants of health

2. Build an evidence base for the field of multi-sector data integration to improve health

3. Utilize the power of peer learning and collaboration
Purpose

1. Share how communities are using mapping and data visualization tools to engage partners in multisector data sharing collaborations

2. Discuss why data visualization is an important tool, how to get started, and key considerations

3. Explore how data is made available to internal and external partners in a way that ensures protection of personal health information
Data Visualization for Population Health Initiatives

Andrew Beck, MD MPH
Cole Brokamp, PhD
November 14, 2016
Objectives

• State the case for using data to inform population health improvement
• Introduce the development of a platform through which data can be visualized (and shared)
• Demonstrate the prototype of the data sharing platform
Morbidity and Mortality
• Reduce annual infant deaths in Hamilton County by 33%
• Reduce disparity in hospital bed days by 15%, focusing on 2 high risk neighborhoods (Avondale and Price Hill)

Thriving
• Ensure 5 year olds have a ‘healthy mind and body’*
• Increase percent of children reading proficiently by 3rd grade in Cincinnati Public Schools from 78% to 90%

*Bundle measure: immunization, BMI, dental, behavior, vision, emergent literacy, speech, hearing
Population health disparities – asthma

- 1,000 in-county children admitted annually
  - Within 12 months:
    - 20% readmitted
    - 40% revisit ED
  - Neighborhood disparities
- Avondale admission rate:
  - >3 times county average
  - >7 times national average
Population health disparities beyond asthma
Disparities in total bed-days across County

Neighborhood Average Annual Inpatient Days, 2012-2014, per 1,000 population ages 0-18 by Hamilton County neighborhood (Excludes patients whose address is at JFS)
Enhance place-based risk assessments

Virtual home or neighborhood visit

- No car
- Mental health challenges
- Mold present
- Food insecure

How can we get to this point?
Current use cases

- Using maps to develop **common language** between health care system and community partners, including:
  - Community pharmacies
  - Housing experts
  - Schools
  - Community organizers
- Prototypes ...
Moving from prototype to platform
App Aims

• Community member versus “expert/researcher”
• Who are you trying to reach?
• Single use / exploratory versus batch processing / statistical
• Make raw data available? downloadable?
Tools

• Open source
• Data fetch/aggregate: bash
• Computation (Amazon EC2/S3, internal computational cluster)
• Geography tools: GEOS, GDAL, PROJ4 (and their respective wrappers for R, Python, qGIS)
• R Shiny (Server) / flexdashboard
• R/Javascript Interface for Interactive HTTP widgets:
  • mapping: Leaflet
  • visualization: Plotly
  • tables: DataTables
Private Health Information Concerns

- Hosting servers (Amazon BAA, internal hosting, HIPPA compliance)
- Geocoding internally
- Jitter location points
- transform to heatmap / krigging
- restricted zooming on interactive maps and images
GIS is more than mapping...

• Interactive and linked visualizations
• Statistics and distributions to put a place on a relative scale
• Predictive modeling and inference with different assumptions or model parameters
• Describing risk has been done, what can we do to manage risk?
Be a part of a reproducible research pipeline

- Script *everything*: download, aggregating, spot fixing, converting projections, merging geospatial objects, saving, exporting
- Allow others (even your future self) to verify and build on your work
- Future fixes and changes are easier
- Work more shareable, easier to collaborate
- Work more scalable when it comes to deployment
- Using text/command based software also allows version management
Application Feedback Cycle

- Track simple visit rate
- Focus groups (users/clinicians/community leaders)
- User studies (click rate, user behavior)
- A/B feature testing
Future Work

• Embed in Electronic Health Record system
• GPS use for current location
• Better co-visualization strategies
• Implement near real time updates for acute exposures and outcomes (air pollution, combined sewer overflow events, allergens, heat, hospital visits and admissions, etc.)
### Data

<table>
<thead>
<tr>
<th>Database</th>
<th>Example Data</th>
<th>Space</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topologically Integrated Geographic Encoding and Referencing (TIGER)</td>
<td>Proximity to roadway, locations of airports, railroads, bodies of water, elevation</td>
<td>Exact location (nationwide)</td>
<td>Yearly</td>
</tr>
<tr>
<td>US Census American Community Survey (ACS)</td>
<td>Population, socioeconomic indicators</td>
<td>Census tract or block (nationwide)</td>
<td>Yearly</td>
</tr>
<tr>
<td>Spatiotemporal PM2.5 Predictions</td>
<td>Estimated air pollution exposure</td>
<td>Exact location (seven county area served by CCHMC)</td>
<td>Daily (2000 – 2015)</td>
</tr>
<tr>
<td>Moderate Resolution Imaging Spectroradiometer (MODIS)</td>
<td>Greenspace, land cover, aerosol optical thickness, surface reflectance</td>
<td>3km square grids (nationwide)</td>
<td>Daily (2000 – 2015)</td>
</tr>
<tr>
<td>Open Street Map (OSM)</td>
<td>Public transportation routes</td>
<td>Exact location</td>
<td>Current</td>
</tr>
<tr>
<td>Global Historical Climatology Network (GHCN)</td>
<td>Temperature, relative humidity, precipitation, wind speed and direction</td>
<td>Interpreted from &gt; 5000 stations (nationwide)</td>
<td>Daily</td>
</tr>
<tr>
<td>National Emissions Inventory (NEI)</td>
<td>Location and amount of emissions</td>
<td>Exact location</td>
<td>Yearly</td>
</tr>
<tr>
<td>EPA Air Quality System (AQS)</td>
<td>Fixed site monitoring of ambient pollutant levels, air quality alerts</td>
<td>Interpreted from over 2500 stations (nationwide)</td>
<td>Daily</td>
</tr>
<tr>
<td>Smart Location Database (SLD)</td>
<td>Location efficiency, characteristics of the built environment</td>
<td>Census tract</td>
<td>2010</td>
</tr>
<tr>
<td>Cincinnati Area Geographic Information Systems (CAGIS)</td>
<td>School districts, crime density, combined sewer overflow sites, housing infractions, etc</td>
<td>Exact location (Greater Cincinnati Area)</td>
<td>Current</td>
</tr>
<tr>
<td>US Department of Transportation (DOT)</td>
<td>Highway traffic intensity</td>
<td>Exact location (nationwide)</td>
<td>Yearly</td>
</tr>
<tr>
<td>National Park Service Geospatial Sound Modeling (GSM)</td>
<td>Noise level</td>
<td>270 m sq grids (nationwide)</td>
<td>2015</td>
</tr>
</tbody>
</table>
Individual dashboard
Comments? Questions?

Andrew.Beck1@cchmc.org
Cole.Brokamp@cchmc.org
Appendix
Application Deployment

- **Docker**: a software containerization platform
- Makes your software stack reproducible from the ground up
- No more software dependency issues (“But it worked on my computer!”)
- Like a VM, but runs on linux kernel
- Makes scaling development
- GUI available for linux/windows/mac to share private images

**shiny_docker**: A robust method to automatically dockerize your R Shiny Application

(https://github.com/cole-brokamp/shiny_docker)

**automagic**: Parse R code to automagically install required R packages ahead of time

(https://github.com/cole-brokamp/automagic)
Allegheny County Data Sharing Alliance for Health (ACDSAH)

Data Visualization Webinar

Vision: a connected data warehouse that provides multi-source data for cross sector decision making to impact the health of the 130 municipalities and 1.2 million residents in Allegheny County.
**Goal:** Build a connected data system of multi-source datasets to improve the cardiovascular health of residents in Allegheny County

- **Data System “Social Determinant Data Warehouse”**
  - Select datasets will be used in a simulated model (FRED model - Framework for Reconstructing Epidemiological Dynamics @ U. Pitt) to examine the distribution of CVD throughout Allegheny County and test interventions in high need areas.

- The data collected can be used on an ongoing basis to support ACHD projects, specifically the Plan for a Healthier Allegheny (PHA)

- Provide both the project partners and citizens with public access to the collection of relevant health determinates datasets
Who We Are

• **Lead Agency** - Allegheny County Health Department

• **ACDSAH partners:**
  - Allegheny County Department of Human Services
  - Allegheny County Economic Development
  - UPMC Health Plan, Gateway Health Plan, Highmark Health Plan
  - Jewish Health Care Foundation and
  - Carnegie Mellon Traffic21 Institute,
  - The University of Pittsburgh GSPH
  - RAND Corporation
The Data

Health Inputs
- Obesity rates
- Smoking rates
- Medical claims data
  - Hypertension
  - Diabetes
  - Hyperlipidemia
  **Diagnosed & Diagnosed + Meds**
  - Co-morbidity
    - Hypertension + Diabetes+
    - Hyperlipidemia (diagnosed)
    - Anxiety medication
    - Depression medication
- Hospitalization
  **(PHC4 zip codes only)**
  - due to diabetes
  - due to hypertension or stroke

Natural Environment
- Air Quality
  - TRI
  - Diesel particulate
- Land Use
  - Woodlands/ forest
  - Greenways Barren Land

Social
- Demographics
  - Age
  - Race
  - Gender
  - Median income
  - Poverty rates
  - Employment Rates
  - Educational attainment
- Access to Transportation
  - Vehicle Ownership
  - Commute time to work
- Homicide
- Age of Death

Built Environment
- Land use
  - Roadways
  - Parks
  - Trails
  - Agriculture land
  - Urban
- Traffic Data
  - 911 response time
  - Hourly Traffic Counts
- Health facilities
  - Primary Care
  - Hospitals
- Vacant properties
- Home ownership/ rentals
- Age of housing
- Walk Scores
- Illegal dump Sites
- Food Access
  - Fast food
  - Farmers markets
  - Supermarkets
- Food deserts
- Tobacco vendors
- Alcohol vendors
- Exempt clean air vendors
Compiling the Data

- Health Department utilizes an internal server as a data repository to store and organize the DASH data. (ArcGIS and Excel are used for geocoding and aggregation)

- Building the data set at the census tract level

- Human Services has developed a data warehouse integrating data sources at the individual level
Data Tools & Features

- Current Visualization tools
  - ArcGIS
    - To identify areas to target

- Online Community Profiles (University of Pittsburgh, University Center Social & Urban Research (UCSUR))
  - To provide community with maps of their data
  - To store data for future analysis

- Healthy Communities Institute (HCI) “community indicators”
  - Public dashboards of county data

- FRED
  - Visual models that demonstrate response to interventions
**Audience:** Public, Staff, Foundations, Health Care, Other County and City institutions

**Pros:** layering of issues, see patterns

**Cons:** not interactive
UCSUR maintains a western PA regional data center where Community profiles allow users to find and download datasets to map and visualize data. (Partnership is facilitated by Allegheny County Statistics)

https://profiles.ucsur.pitt.edu/profiles/census-tract/42003202300/education/

**Audience:** researchers, community organizations; county departments, foundations

**Pros:** gives interactive ability to layer maps, currently available at no cost

**Cons:** too academic
**Audience:** public, ACHD, community organizations

**Pros:** very user friendly, easy to understand, continually updated

**Cons:** requires work to input new data, annual licensing fee, no input into updates

http://www.achd.net/aci/index.html?hcn=%2Findex.php%3Fmodule%3Dindicators%26controller%3Dindex%26action%3Dview%26indicatorId%3D2064%26localeId%3D2297%26hcnembedredirect_3%3D1
“FRED”

- Uses a modeling system based on synthetic populations that capture the demographic and geographic distributions of the population.

- A strength of the model is its ability to visualize patterns of incident occurrence and predict how interventions may change the way a disease looks.

- FRED has the potential for live action visualization.
- Tableau (ACHD)
  - for more interactive dashboards
  
  - **Pros:** interactive, ability to download data directly, data can be shared with other jurisdictions
  
  - **Cons:** annual cost, training curve
Mental Health Services Provided by DHS

- **66,816** Clients Received MH Services
- **$231,843,242** were Paid for MH Services

**FY Comparison**

- **Clients Served**: FY15/16 > FY16/17
- **Amount Paid**: $0M > $200M

**Trend by Fiscal Quarter**

- Q1: $57,062,694
- Q2: $57,010,065
- Q3: $59,291,923
- Q4: $57,678,861

**Clients by Gender, Age and Race**

- White, Black, Other, Unknown

**By Service Category**

- OP-MH: 53%
- Admin Mgmt: 20%
- Serv Coord: 17%
- Crisis: 14%
- Medi Check: 39%
- BHR S$: 8%
- IP-MH: 7%
- Consults: F/S

**By Diagnosis Category**

- Major Depression: 29%
- Adjustment Disorders: 12%
- Anxiety Disorders: 11%
- Bipolar Disorder: 13%
- Other: 2%

Please note that there is a time lag between the date of service and when the claim is submitted. This dashboard only includes MH claims submitted before 5/16/2016.

Source: https://tableau.alleghenycounty.us/#/signin?externalRedirect=%2Fviews%2FMHDashboardforDHS%2FMHDashboardforDHS%3F:embed%3Dy&site=
Lessons Learned

• No one visualization strategy meets all needs
  – Academics, open data, community members

• Try to build on existing opportunities for sustainability

• Is it best to use many or choose one?
Next Steps

- Work internally with IT to track the types of users & downloads that are currently happening on the county website.

- Identify the audience ACHD is reaching and select the next visualization strategies with the goal of reaching a broader audience.

- Choose a visualization strategy(s) that allows both more in-depth data downloads for organizational reports & partner grant writing in addition to providing the general public with the information they desire.
PRIMARY AIMS

Minimize asthma outside clinic walls

Children age 2-12

Households making < 80% of AMI
HHDSM SO FAR

ETHNICITY
- Hispanic 24%
- Black 29%
- White 26%
- Sudanese 3%
- Biracial 3%
- Asian 9%
- Bosnian 3%

HEALTH INSURANCE COVERAGE
- Medicaid 85%
- Private 6%
- Hawk-i 9%

GENDER
- Boys 71%
- Girls 29%

AVERAGE AGE
The average age of children referred to HHDSM increased slightly over the last quarter from 6 to 7 years.
DATA ANALYSIS
DATA SHARING
DATA VISUALIZATION

WHAT IS THE BURDEN OF ASTHMA?

$1.3 million

AMOUNT MEDICAID PAID FOR ASTHMA RELATED CLAIMS OF CHILDREN IN POLK COUNTY

CHARACTERISTICS OF HHDSM CHILDREN

29% GIRLS

71% BOYS

average age 7

AVERAGE REPAIRS AND COSTS

$7,858

$5,642

$4,247
INTENDED AUDIENCES

Public Users

Internal Users
CHALLENGES EXPERIENCED + LEARNINGS
Claire Richmond
Healthy Homes Des Moines
CRICHMOND@PCHTF.ORG
515 232 3233

Brett Burkhart
Shift Interactive
515 221 8645
BRETT@INTERACTIVESHIFT.COM
Connect with Us!

- Sign up for news from All In at dashconnect.org
- Follow us at @DASH_connect and @AcademyHealth at #CHPHealthIT
- Contact information for speakers
  - Andrew Beck - Andrew.Beck1@cchmc.org
  - Cole Brokamp - Cole.Brokamp@cchmc.org
  - Karen Hacker - Karen.Hacker@AlleghenyCounty.US
  - Eric Hulsey - Eric.Hulsey@AlleghenyCounty.US
  - Claire Richmond - CRichmond@pchtfc.org
  - Brett Burkhart - brett@interactiveshift.com

- Evaluation

- A resource list, slides, and recording will be available