Geospatial Analysis of the Electronic Health Record

Jim Gay

- Information Analyst III
- Enterprise Analytics & Reporting Team
- Experience: 5 years (CHOP), 10 years (Healthcare), 18 years (BI/Data Warehouse)

- **Children's Hospital of Philadelphia (CHOP)**
  - The Children's Hospital of Philadelphia is consistently ranked one of the top pediatric hospitals in the nation in the *U.S. News & World Report's* Honor Roll of the Best Children's Hospitals
  - Largest pediatric network in the United States with over 50 locations in Pennsylvania and New Jersey.
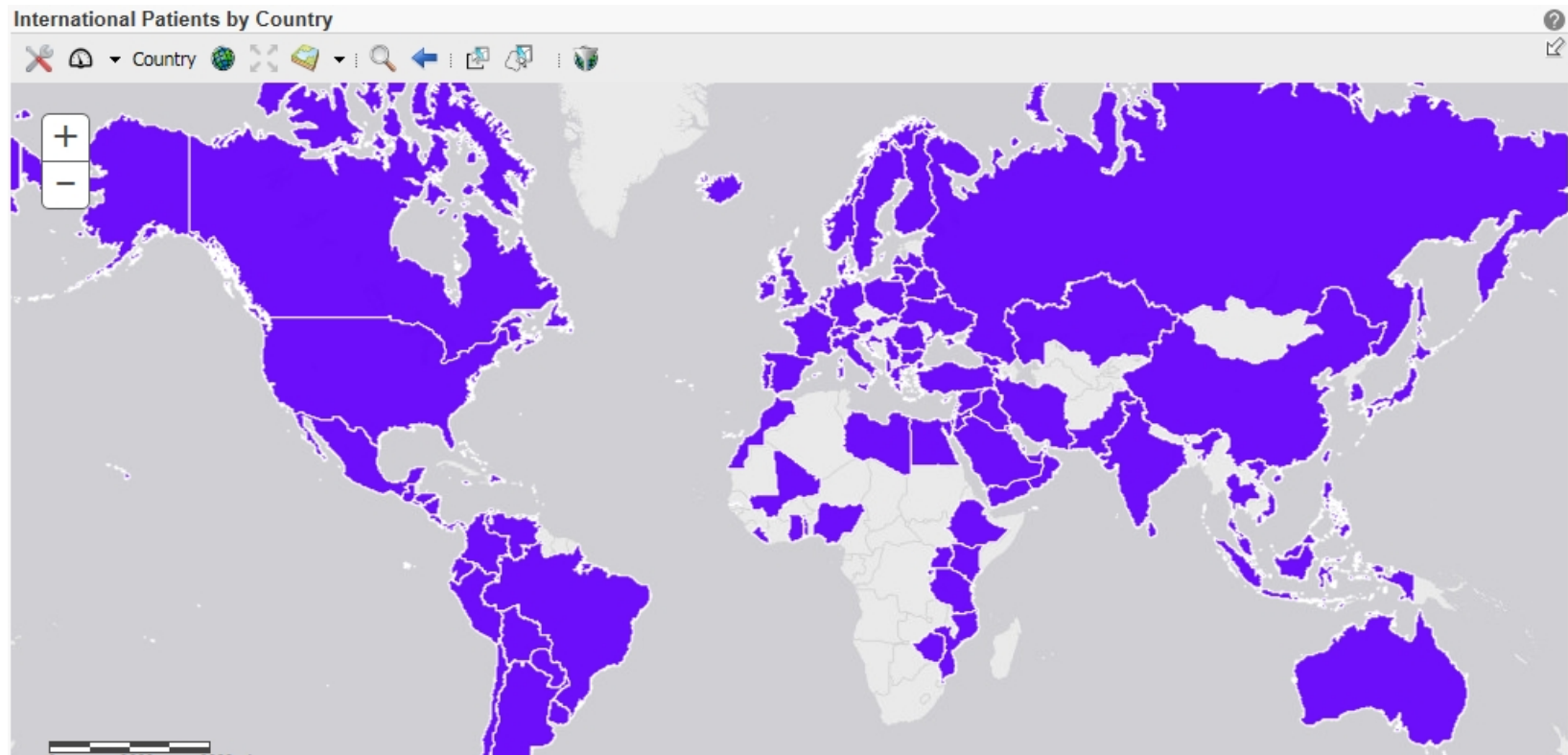 - CHOP's Information Services department named a Top 25 ELITE 100 organization.

Agenda:

- **Geocoding our patient population**

- **Review the automated geocoding process**

- **Review the results**

- **Integrating GIS into our health record**

- **Questions?**

CHOP's Worldwide Footprint

Patients visit CHOP from over 110 countries

International Patients by Country

Our infrastructure

- EPIC electronic health record

- Patient population of 2.5 million (6 million addresses in total)

- Employee population of 15,000

Filter data for geocoding

- **The following addresses were omitted from geocoding:**
  - Missing address line one
  - Missing zip code
  - International patients
  - Test patients

Address Data Quality

- Over 300 different spellings of 'Philadelphia' in our EHR.

- Examples:

|		
PHILADELPHIA/09	PHILADELPA	PHILADEIB15
PHILADELPHIAA	PHILADELPGIA	PHILADEILPHIA.
PHILADELPHIAL	PHILADELPH	PHILADEKPHIA
PHILADELPHIALS	PHILADELPHA	Philadel
PHILADELPHIAP	PHILADELPHAI	Philadel[phia
PHILADELPHIAY	Philadelphai	PHILADEL=HIA
PHILADELPHIIA	PHILADELPHAP	PHILADEL=PHI
PHILADELPHIL	PHILADELPHEA	Philadelaphia
PHILADELPHILA	PHILADELPHGIA	PHILADELDELPHIA
Philadelphila	philadelphi	PHILADELEPHIA

Geocoding Process – Overview

- Import filtered addresses (patients and employees)

- Geocode the locations using ESRI 2014 USA Geocoding Service

- Spatial join to identify the census tract for each patient and employee

- Store information in our geodatabase (latitude, longitude, metadata, census tract)

Geocoding Process – Import & Geocoding
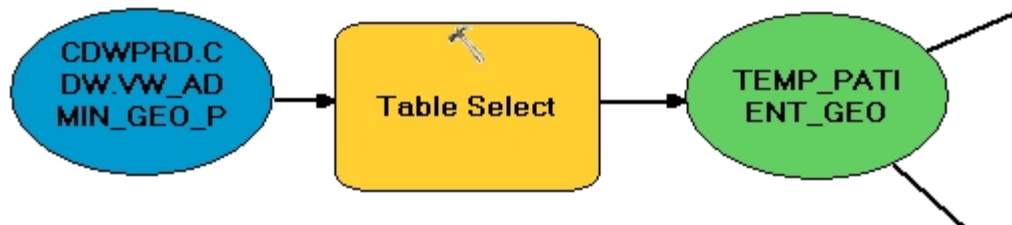
Import new patients & employees along with updated addresses
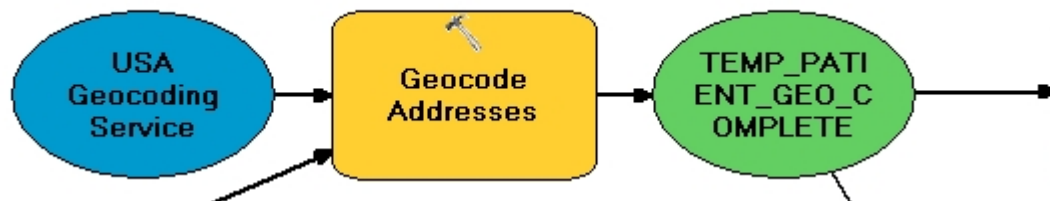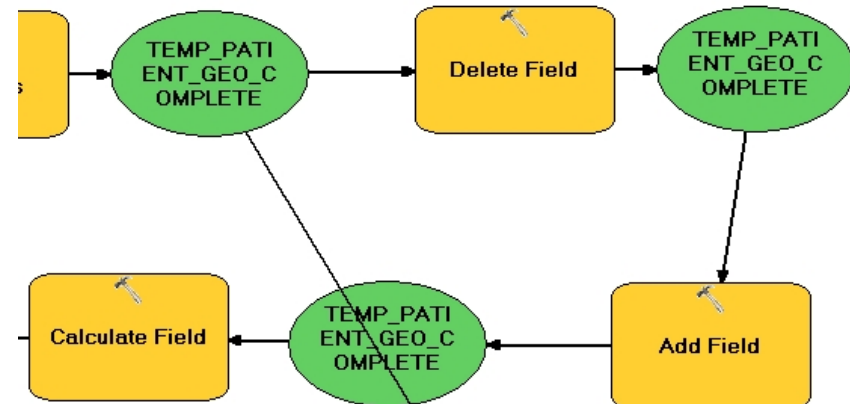

Geocode using ESRI 2014 USA Geocoding Service

Geocoding Process – Data Cleanup

- **Remove fields**
  - Match Type
  - ARC_Address
  - ARC_City
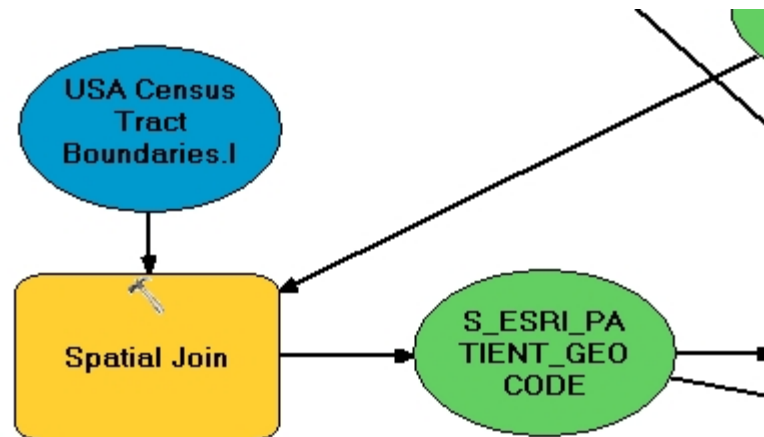 - ARC_State
  - ARC_Zip
  - Original address fields from EHR

- **Add field**
  - Add field GEOCODING_SOURCE to dataset

Geocoding Process – Spatial Join

- Identify census tract for all patients and employees

- Source is USA Census Tract Boundaries

- Most time consuming part of the process

Geocoding Process – Database Export

- Before exporting extra fields are removed from the dataset that were part of the spatial join

- 3 stage tables are exported to our data warehouse
  - PATIENT_GEO
  - PATIENT_GEO_ADDRESS_HISTORY
  - EMPLOYEE_GEO
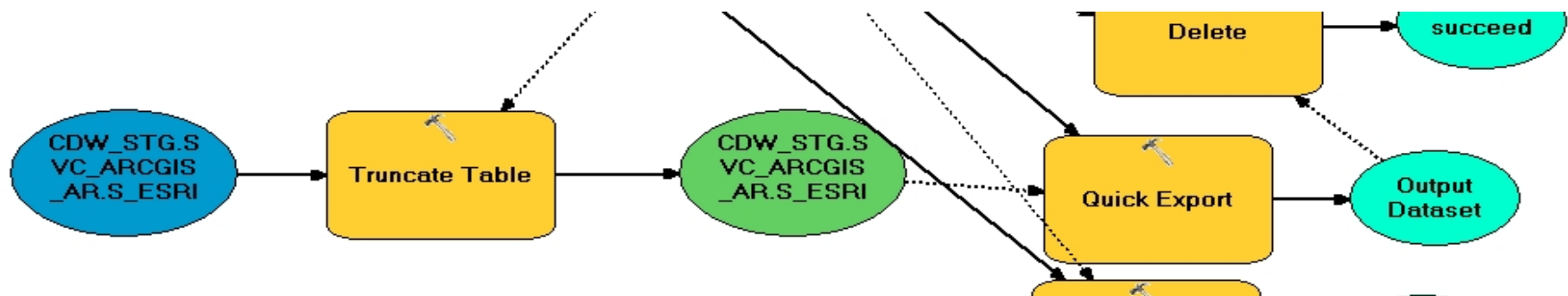
Geocoding Process – Final Dataset

FIPS	34007602902
STCOFIPS	34007
SPAT_STATUS_NM	MATCH
LOCATOR_NM	ADDRESS_POINTS
ACCURACY_SCORE	97
STREET_LONG_DEG_X	-75.22337765
STREET_LAT_DEG_Y	39.33111701
ROOFTOP_LONG_DEG_X	-75.321456
ROOFTOP_LAT_DEG_Y	39.369258
SRC_ADDRESS	3501 CIVIC CENTER BLVD, PHILADELPHIA, PA, 19104
STREET_SIDE	RIGHT
GEOCODING_SRC	2014 USA GEOCODING SERVICE

Geocoding Process – Python Scripts

- **Geocoding model exported to python scripts (patient, patient address history, and employee)**

- **Python scripts are scheduled using a .BAT file on Windows Server**

- **An email is sent out once the process is complete each morning**

- **Processing time for the 3 scripts is about 7 minutes each day**

Geocoding Process – Results

- **To date, 6.7 million addresses have been geocoded using this process for both patients and employees**

Geocoding Status	# of Addresses	% of Total	Accuracy Score
Match	5.85 mil	86.8%	97.4
Tie	825k	12.2%	82.1
Unmatched	62k	0.9%	0

Address Locator	% of Total	Accuracy Score
Address Points	77.9%	95.9
Street Address	14.6%	91.4
Zip Code	7.4%	100
City State	0.1%	95.1

Geocoding Improvements

- **Adding Census Block to our geodatabase**
  - We determined that census tract was not giving us the level of detail we needed so we are setting up a process to add census block to our geodatabase

- **Potential improvements**
  - Reverse geocoding
  - Address clean up
  - Eliminate free text for city and state within our electronic health record
  - Only one geocoding failure in the past 13 months

Geocoding Failure

Address Line 1:

{"ATTACHMENTS":[{"__TYPE":"ITEMIDATTACHMENT:#EXCHANGE","ITEMID":{"__TYPE":"ITEMID:#EXCHANGE","ID":"AAMKADY0MJZINGNHLTEYZWMTNDI4NI04ZTLMLTRHYTJIMDQXODIXYWBGAAAAAACHIU8BRGY3T6P6EDSKOCXNBWCZ4IWRXUCTQIKRGSPY+MEJADH0FX1TAABRFDMGBEXTSZXPFMMW8AZKAAE11VEVAAA=","CHANGEKEY":NULL},"NAME":"LIMITED TIME! SUMMER FLASH SALE = 40% OFF EVERYTHING + EXTRA 15% OFF","ISINLINE":FALSE}]}

GIS @ CHOP – Services Implemented

- **ArcGIS for Server**

  - Feature services for boundaries implemented (Countries, States, Counties, Zip Codes, and Census Tracts).

- **ArcGIS for Desktop**

  - Used by our analytics team and strategy team for various GIS analysis.
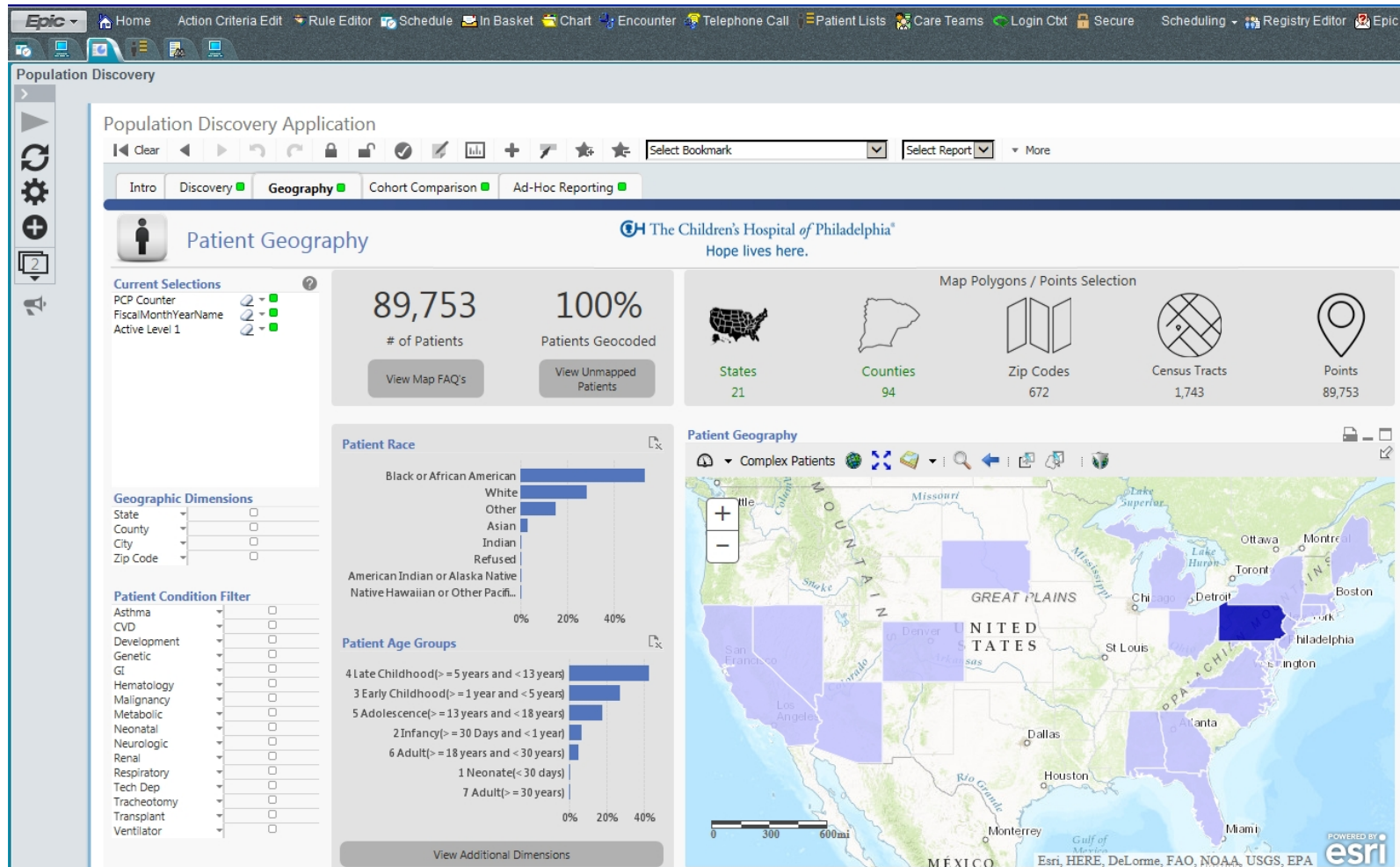
- **Business Analyst**

  - Used for geocoding, routing, and market analysis.

- **KliqMaps extension for Qlikview**

  - Plugin that runs within our Qlikview applications that integrates our analytic tools with our ESRI services.
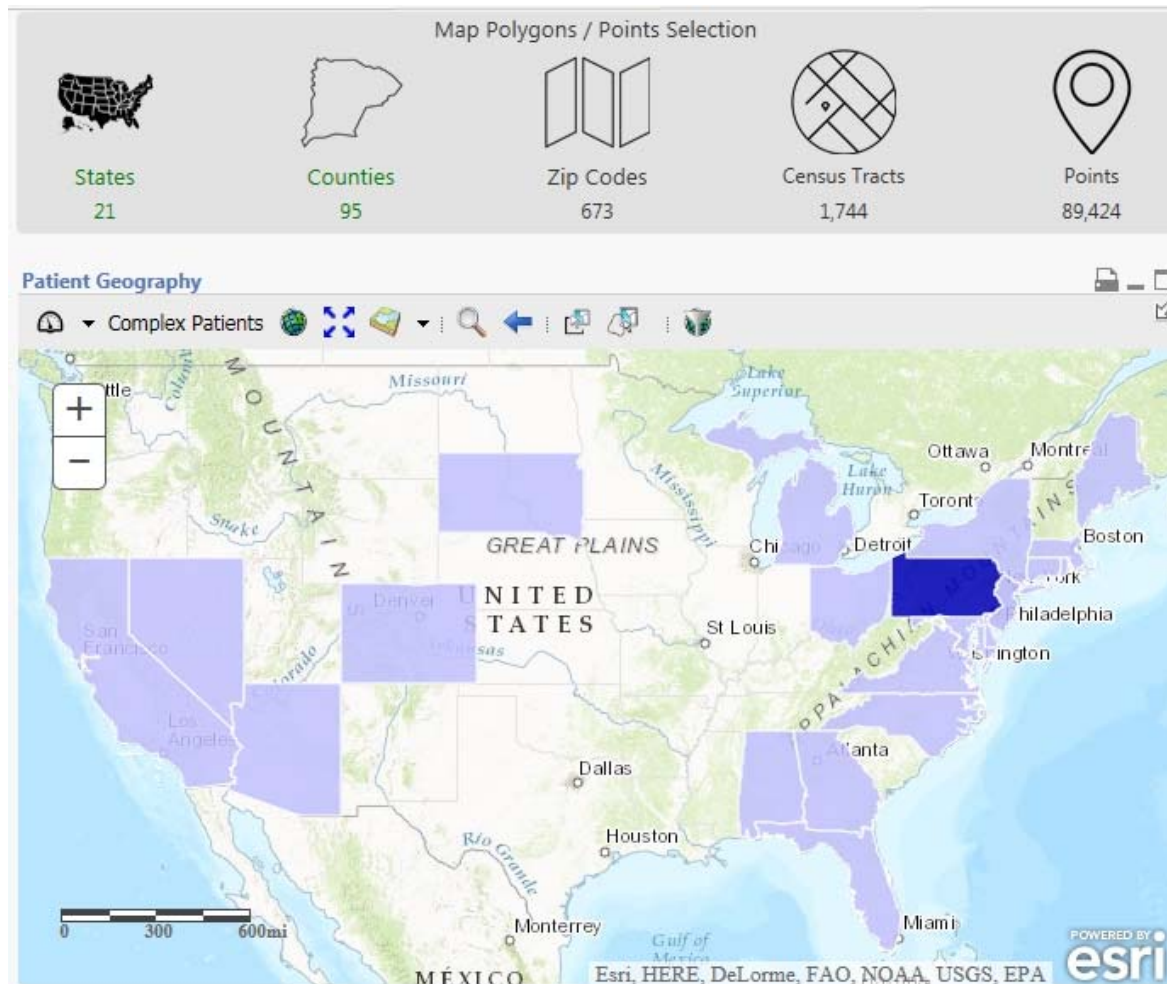
Integrating GIS into our EHR

Geospatial Analysis with Qlikview

Geospatial Analysis @ CHOP

- **Information Services department – Alternative work arrangement (AWA)**
  - A study was done to determine how far each IS employee traveled to work each day.
  - The original policy was that each IS employee could work from home 1 day per week.
  - After the analysis was completed a change was made to the IS policy allowing employees to work from home 2 days per week.

- **Planning, Strategy, and Business Development**

- **Asthma**

- **Population Health**

Questions ?

Contact:

Jim Gay
gayj@chop.edu