Exploring Data Opportunities at CPHIT: MHCC and OptumLabs Data

Dr. Jonathan Weiner
Elyse Lasser
Thomas Richards
OVERVIEW

• CPHIT
  • Brief overview
  • Research priorities

• Maryland Healthcare Cost Commission
  • Brief overview
  • Data availability
  • Process to access data

• OptumLabs
  • OptumLabs/CPHIT arrangement
  • Available Data
  • Next Steps
CENTER FOR POPULATION HEALTH INFORMATION TECHNOLOGY
Mission: To improve the health and well-being of populations by advancing the state-of-the-art of Health IT across public and private health organizations.

Focus: The application of electronic health records (EHRs), “big data” and other e-health tools targeted at communities and populations.
Research Portfolio *(selected list)*

- **eACG**: Developing a wide range of state of the art EHR-based population focused predictive modeling “e-ACG” tools
  - EHR ICD and e-prescribing Rx input (coordinated with claims)
  - Non-claims data: Lab Data; BMI and Vitals; and Social History

- **VHA**: Collaborating with the Veteran Health Administration
  - Developed population health analytic framework
  - Linking geo and social data for obesity trend analysis
  - Adding social/geo framework to PCMH (PACT) case finding program

- **Frailty**: Developing a new geriatric/frailty “e-risk” score utilizing structured and unstructured EHR data and Claims
- **Falls**: Developing analytics with Baltimore City Health Department for regional collaboration to identify and predict elder’s fall injuries in the community using social, medical, and public health data.

- **Opioid**: Using multiple novel sources of data to develop predictive models to identify persons at risk for opioid overdose.
  - CRISP/HIE’s PDMP data; Department of Justice; & DHMH

- **Consumer Data**: Linking consumer/marketing data with medical data to identify health outcomes (e.g., consumer reports).

- **Pop e-Measures**: Collaborating with Maryland State Health Department, HIE and Hospital Commission.
  - Assist in the building of a statewide population health digital measurement infrastructure for “waiver” and beyond.
  - Develop population health focused measures.
SELECT LIST OF DATA SOURCES

- Sample list of data sources (accessed or hosted; fully or partially):
  - IMS Claims
  - Optum Labs Claims + EHR
  - Claims and EHR records of various national and regional integrated health delivery systems
    - VHA, Atrius, HealthPartners
  - Maryland’s HSCRC (case-mix / hospital discharges)
  - Maryland’s MHCC (all-payer claims database)
  - Maryland’s HIE (CRISP; health information exchange)
  - and other novel sources of data…
Sample list of data sources (accessed or hosted; fully or partially):

- IMS Claims
- **Optum Labs Claims + EHR**
  - Claims and EHR records of various national and regional integrated health delivery systems
    - VHA, Atrius, HealthPartners
  - Maryland’s HSCRC (case-mix / hospital discharges)
- **Maryland’s MHCC (all-payer claims database)**
  - Maryland’s HIE (CRISP; health information exchange)
  - and other novel sources of data…
MARYLAND HEALTH CARE COMMISSION

Maryland Medical Care Data Base (MCDB or APDB)
• Brief overview
• Data availability
• Process to access data
MHCC: ALL-PAYER CLAIMS DATABASE (OVERVIEW)

• An independent regulatory agency whose mission is to plan for health system needs, promote informed decision making, increase accountability, and improve access in a rapidly changing health care environment (http://mhcc.maryland.gov/)

• Made up of 4 centers:
  • Center for Health Care Facilities Planning and Development
  • Center for Health Information Technology and Innovative Care Delivery
  • Center for Analysis and Information Systems
  • Center for Quality Measurement and Reporting
• Previously, each request had to go directly through MHCC and commissioner review board

• JHU has entered into an agreement with MHCC to receive the MCDB at no cost for 2 years for JHSPH faculty to perform pilot projects.
  • Faculty/students will need to apply internally for approval
MHCC: ALL-PAYER CLAIMS DATABASE: DATA

- Enrollment, provider, and claims data for Maryland Residents enrolled in private insurance (some medicare fee-for service, Medicare Advantage, Medicaid Managed care is in the database, but not approved for data release) Does not include self-insured ERISA health plans

- Years available: 2010-2015 with data refresh for 2016 when available

- Includes Dates of Service, Date of Birth, and 5 digit Zipcode

- Standard file includes 4 files:
  - Eligibility information
  - Professional services claims
  - Institutional claims
  - Pharmacy claims

- Data dictionary and Application for data request can be found in JHBox: http://preview.tinyurl.com/cphitmhcc

<table>
<thead>
<tr>
<th></th>
<th>Market</th>
<th></th>
<th>Large Employers</th>
<th>Small Employers</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Members</td>
<td>Total members as of December 31</td>
<td>1,229,516</td>
<td>772,714</td>
<td>213,075</td>
<td>243,727</td>
</tr>
<tr>
<td></td>
<td>Total member months</td>
<td>14,993,456</td>
<td>9,361,152</td>
<td>2,634,609</td>
<td>2,997,695</td>
</tr>
<tr>
<td>Spending</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMPM spending, all services combined</td>
<td></td>
<td>$376</td>
<td>$365</td>
<td>$368</td>
<td>$411</td>
</tr>
<tr>
<td>PMPM OOP, all services combined</td>
<td></td>
<td>$71</td>
<td>$55</td>
<td>$87</td>
<td>$105</td>
</tr>
<tr>
<td>PMPM OOP, Medical Only</td>
<td></td>
<td>$55</td>
<td>$42</td>
<td>$66</td>
<td>$87</td>
</tr>
<tr>
<td>PMPM OOP, Prescription Drugs</td>
<td></td>
<td>$16</td>
<td>$13</td>
<td>$22</td>
<td>$18</td>
</tr>
<tr>
<td>PMPM Spending By Service Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inpatient Hospital Facility</td>
<td></td>
<td>$61</td>
<td>$63</td>
<td>$55</td>
<td>$62</td>
</tr>
<tr>
<td>Inpatient Non-Hospital Facility</td>
<td></td>
<td>$0.5</td>
<td>$0.5</td>
<td>$0.4</td>
<td>$0.8</td>
</tr>
<tr>
<td>Outpatient Hospital Facility</td>
<td></td>
<td>$71</td>
<td>$68</td>
<td>$67</td>
<td>$86</td>
</tr>
<tr>
<td>Outpatient Non-Hospital Facility</td>
<td></td>
<td>$8</td>
<td>$8</td>
<td>$8</td>
<td>$10</td>
</tr>
<tr>
<td>Professional Services</td>
<td></td>
<td>$109</td>
<td>$107</td>
<td>$108</td>
<td>$115</td>
</tr>
<tr>
<td>Labs/Imaging</td>
<td></td>
<td>$32</td>
<td>$29</td>
<td>$32</td>
<td>$39</td>
</tr>
<tr>
<td>SubTotal (Medical Only)</td>
<td></td>
<td>$282</td>
<td>$276</td>
<td>$270</td>
<td>$313</td>
</tr>
<tr>
<td>Prescription Drugs¹</td>
<td></td>
<td>$94</td>
<td>$90</td>
<td>$98</td>
<td>$98</td>
</tr>
</tbody>
</table>

Note:  
(1) Prescription drug spending is missing for some Federal Employees Health Benefits Program (FEHBP) members; therefore, drug spending for large employers is calculated using only members who have the prescription drug benefit.  
(2) FEHBP members are included in the large employer market.  
(3) PMPM portion of spending for insurers is overall PMPM (all services combined) less PMPM OOP (all services combined).  

Analyses using MHCC’s MCDB for 2015  
Source: MHCC 8/17 report
MHCC: ALL-PAYER CLAIMS DATABASE: COUNTS 2010-2014

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>7,476,648</td>
</tr>
<tr>
<td>Professional encounters</td>
<td>370,557,205</td>
</tr>
<tr>
<td>with non-null zip</td>
<td>370,524,870 (99.99%)</td>
</tr>
<tr>
<td>Institutional encounters</td>
<td>17,288,503</td>
</tr>
<tr>
<td>with non-null zip</td>
<td>17,286,606 (99.99%)</td>
</tr>
<tr>
<td>Pharmacy claims</td>
<td>102,035,561</td>
</tr>
<tr>
<td>Diagnosis lines</td>
<td>767,286,922</td>
</tr>
<tr>
<td>Procedure lines</td>
<td>403,417,139</td>
</tr>
</tbody>
</table>

* 2015 is in house and is currently going through the intake process
MHCC: ALL-PAYER CLAIMS DATABASE: PATIENT OVERLAP
MHCC: ALL-PAYER CLAIMS DATABASE: TRANSFORMATION PIPELINE

Raw Data → Extract and Transform → CPHIT Research Data Model → Data Mining → Analytic Files

Claims & Eligibility
MHCC: ALL-PAYER CLAIMS DATABASE: CPHIT RESEARCH
DATA MODEL
### Encounters

1. **3/1/2017**  
   - Delivery  
   - Obesity

2. **7/1/2017**  
   - IUD Insertion  
   - Hypertension  
   - Obesity  
   - Asthma

3. **9/1/2017**  
   - IUD Surveillance

4. **12/1/2017**  
   - IUD Removal

### Monthly Events

- Delivery 03  
- OBESITY 03  
- IUD INSERTION 07  
- HTN 07  
- OBESITY 07  
- ASTHMA 07  
- IUD SURVEILLANCE 09  
- IUD REMOVAL 12

### Event Vectors

- DELIVERY 001000000000  
- OBESITY 001000100000  
- IUD INSERTION 000000100000  
- HTN 000000100000  
- ASTHMA 000000100000  
- IUD SURVEILLANCE 000000010000  
- IUD REMOVAL 000000000001  
- PREGNANT 111000000000  
- POSTPARTUM 000110000000  
- IUD INSERTED 000000111111
MHCC: ALL-PAYER CLAIMS DATABASE: APPLICATION PROCESS

• Review Data Dictionary and complete application found in JHBox

• Receive IRB approval on project (only IRB approved projects can access MHCC data)

• Set up SAFE Virtual Desktop with CCDA for data access
  • Potential cost from IT@JH for storage: $4 per 100GB per month
    • Depends on the extracted data users need
  • Data must remain at the Mount Washington Data Center

• CPHIT can assist in:
  • Developing full analytic data set
  • Provide small sample of data for researcher to write code to pull full data including inclusion/exclusion criteria.
  • Involvement will be dependent on support needed
OPTUMLABS
Unified View (linked EHR and Claims data)
Who is OptumLabs?

An entity of United Health Group

- United Healthcare provides health benefits to 45 million people in the US and Brazil
- Optum provides Health services to 5,000 hospitals and 80,000 physician practices and other health care facilities

OptumLabs was founded as a partnership between Optum and Mayo Clinic as an open center for collaborative research and innovation with a state of the art facility based in Cambridge, MA.
Provide **access to the Unified View** (Linked EHR/Claims database) to 10 faculty at low to no cost based on project funding availability.

Goal is to meet specific **financial targets** based on the number of faculty signed up on the RAP (5 faculty minimum, 10 faculty max)

Funded projects should **budget 20% of direct costs** for OptumLabs data

Ability to collaborate with a number of other partners

All projects are reviewed by their research council

OptumLabs has **regular presentations and ideas exchange meetings** for researchers to share work in progress.
Overview of Optum Labs Databases

Our Data Today: Claims, EHRs & Consumer Behavior

315 million U.S. population

>33 million unlinkable claims

>35 million Consumer

>128 million linkable claims

1,500+ data fields:
- Medical claims
- Pharmacy claims
- Lab claims and results
- Health risk assessments
- Standardized costs of care
- Race
- Income
- Education level
- Household
- Geography
- Mortality

Tests, Treatments

>31 million Clinical

Mayo

Expanded insights with AARP member data

300+ additional data fields:
- Consumer Behavior AARP: general trends
- Demographic view including Income, Assets, Home Value, Education Level, Marital Status, Occupation, Home Ownership, Household Make-Up (multi-generational, presence of: children, grandchildren, grandparents), Ethnicity Data
- Psychographic Data including interest and participation in: travel, various leisure activities, charitable giving, advocacy, volunteering, community involvement

Expanded insights with deeper clinical context

250+ additional data fields:
- Encounters
- Vitals (BMI, BP, Heart Rate ...)
- Labs
- Medication orders
- Procedures
- Admissions, discharges and transfers
- Patient-provided information
OPTULABS: UNIFIED VIEW

- Administrative Claims
  - Enrollment, medical and pharmacy
  - Approximately 128 million individuals
  - Varying duration (as far back as 20 years)
  - Commercially insured, some Medicare Advantage and Medicare PDP

- Electronic Health Record
  - Encounters, Signs, diseases and symptoms, Observations and clinical findings, medication administration
  - Approximately 31 million patients
  - Longitudinal data vary up to 8 years
  - From 30+ Large provider practices and IDNs
  - Tables with extractions from clinical notes
• Overlap is about 10 million patients
• At the Census Region only
• Education/Income data from the EHR only
OPTUMLABS: ACCESS

- Access through VDI (virtual portal) with Vmware Horizon Client
  - Windows Server 2012
  - 2 factor authentication (password and text)
  - Can bring own tools/software, but OptumLabs must install
  - sFTP process to move files in and out (no patient level data)
  - PostgreSQL database
OPTUMLABS: NEXT STEPS

• Identify 5-10 faculty who would like to explore the OptumLabs Unified view

• Determine log-ins (at most 2 per named faculty)

• OptumLabs:
  • Perform a boot camp/onboarding process with all named faculty to
  • Provide limited analytic support for each faculty