Overview

1. The Geisinger Health System and the Geisinger Environmental Health Institute
2. The use of electronic health records (EHR) in environmental epidemiologic research
   a) Common methods
   b) Examples: built environment, animal feeding operations
3. Studies of Marcellus shale & health at Geisinger
   a) Overview, summary, design, specific aims
   b) Efforts in exposure assessment using secondary data and their many complexities and challenges
4. Questions
The Geisinger Health System

- Geisinger Clinic, Geisinger Health Plan, hospitals
- 40+ community practice clinics and 6+ hospitals
- 430,000+ primary care patients representative of general population in region
- 2M+ specialty care patients; 2.6M OPT visits
- Electronic health record (EHR) with > 12 y data
- Across a large, varied geography (44+ counties)
- 30% of primary care patients have Geisinger Health Plan insurance – can get claims data
  - Need not have GHP to use health system
- Environmental Health Institute (EHI) started 2007
Environmental Health Institute
Environmental Epidemiology Studies to Date

<table>
<thead>
<tr>
<th>Health Outcome</th>
<th>Environmental Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marcellus shale ENERGY, LAND, WATER</td>
</tr>
<tr>
<td></td>
<td>Built environment LAND, FOOD, PHYSICAL ACTIVITY ENVIRONMENTS</td>
</tr>
<tr>
<td></td>
<td>Abandoned coal mines ENERGY, LAND, WATER</td>
</tr>
<tr>
<td></td>
<td>Animal feeding operations FOOD, LAND, WATER</td>
</tr>
<tr>
<td></td>
<td>Social environment</td>
</tr>
<tr>
<td>Asthma</td>
<td>✓</td>
</tr>
<tr>
<td>Cardiovascular outcomes</td>
<td>✓</td>
</tr>
<tr>
<td>Chronic rhinosinusitis</td>
<td>✓</td>
</tr>
<tr>
<td>Diabetes</td>
<td>✓</td>
</tr>
<tr>
<td>MRSA</td>
<td>✓</td>
</tr>
<tr>
<td>Obesity, child</td>
<td>✓</td>
</tr>
</tbody>
</table>

NIH = NIH-funded

Methods Common to All Studies

- Obtain patient data from EHR
- Geocode patients
- Consider how environment contributes to disease burden and what secondary data sources may assist with exposure assessment
- Use geographic information systems (GIS) to create exposure metrics
- Link exposure and patient measures
- Biostatistical analysis – person, place, time
Johns Hopkins Systems-oriented Childhood Obesity Center
(U54 HD070725)

Project 1
- 163,820 children 3-18 years in 1,289 communities
- Health data from 2001-12, average of 3 annual BMIs per child

- 3 to 6 years
- 7 to 12 years
- 13 to 18 years

Industrial Animal Operations in Relation to MRSA Infection (thesis research of Joan Casey)

- Obtained Nutrient Management Plans for CAFOs, CAOs, VAOs, for swine & dairy/veal operations

Epidemiol Infect 2013
JAMA Intern Med 2013
Environ Health Perspect 2014
Drilled wells increased from 36 in 2006 to 1957 in 2011.
At full build out, estimated to be 20,000-50,000 wells.

Average Production Decline Curve for Marcellus Shale Gas Well

Yearly Declines:
- First Year = 47%
- Second Year = 66%
- Third Year = 71%
- Fourth Year = 47%
Estimated 170 NG compressor stations (MapSearch, Inc. 2011) and ~540 UNGD compressor stations (DEP 2013)

Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania

Environ. Res. Lett. 8 (2013) 014017 (8pp)

Table 5. Statewide emissions estimates for shale gas development and production in 2011.

<table>
<thead>
<tr>
<th>Activities</th>
<th>VOC</th>
<th>NOₓ</th>
<th>PM₂.₅</th>
<th>PM₁₀</th>
<th>SOₓ</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Transport</td>
<td>31–54</td>
<td>550–1000</td>
<td>16–30</td>
<td>17–30</td>
<td>0.82–1.4</td>
</tr>
<tr>
<td>(2) Well drilling and hydraulic fracturing</td>
<td>260–290</td>
<td>6600–8100</td>
<td>150–220</td>
<td>150–220</td>
<td>6.6–190</td>
</tr>
<tr>
<td>(3) Production</td>
<td>71–1800</td>
<td>810–1000</td>
<td>15.78</td>
<td>15.78</td>
<td>4.5–6.2</td>
</tr>
<tr>
<td>(4) Compressor stations</td>
<td>2200–3800</td>
<td>9300–15000</td>
<td>280–1100</td>
<td>280–1100</td>
<td>0.340</td>
</tr>
<tr>
<td>Total*</td>
<td>2500–11000</td>
<td>17000–28000</td>
<td>460–1400</td>
<td>460–1400</td>
<td>12–540</td>
</tr>
</tbody>
</table>

* These totals are reported to two significant figures, as are all intermediate emissions values in this document. The activity emissions may not exactly sum to the totals.
Geisinger Marcellus & Health Studies

• Investigator-initiated projects – in the EHI
  – Sara Rasmussen & Joan Casey, PhD candidates, JHBSPH. Health studies, model EPA air data in relation to UNGD, & radon.

• System-wide initiative
  – Efforts: Use of EHR data for health studies; sharing of clinical data by multiple PA health systems; and long-term involvement in this industry and health studies.
  – Funding from Degenstein Foundation
    • Exposure assessment activities
    • Projects to date: water quality in 1 county (USGS); focus groups – community attitudes and concerns; traffic & trauma study
Specific Aims (A&P Study)

- **(SA1):** Use geocoded well and infrastructure data to develop individual-level exposure models for residences that account for spatial and temporal variation in well development activities based on density, accessibility, and clustering of well development over time, as surrogates for potential exposure to air and water pollutants associated with development and production.

- **(SA2):** In a cohort analysis, evaluate associations of exposure metrics with pregnancy outcomes [...] compare to Elaine Hill’s work.

- **(SA3):** In a case-control analysis, evaluate associations of the exposure metrics with measures of asthma control, comparing asthma patients with and without selected outcomes.

Primary Health Outcomes in Planned Analyses

**Pregnancies** (funded)
- 22,000+ in GHS from 2006-12
- ~10,000 were primary care patients
- Primary outcomes
  - Apgar score, SGA, LBW, pre-term, pre/eclampsia, gestational diabetes

**Asthma** (funded)
- ~38,000 patients with at least 2 encounters with ICD-9 for asthma
- Primary outcomes
  - Asthma control test, asthma exacerbations (IPT, ED, OCS), asthma medication intensification

**Cardiovascular** (pending)
- With Guthrie Health System
- Thousands of events; excellent market coverage in northern tier counties
- Primary outcomes
  - Hospitalizations for acute myocardial infarction, ischemic stroke, congestive heart failure (N = 7K, 5K, and 15K estimated, respectively)
- Secondary outcomes of PVD, TIA, coronary revascularization procedures
Data Sources for Exposure Assessment

- UNGD wells
  - DEP, Powdermill, PA\textsuperscript{IRIS}
  - Source documents – more on following slides
  - FracFocus
  - SRBC Post-hydrofracture Report and Certification of Fluid Disposal
- Compressor stations for UNGD
  - DEP field offices
  - Scanning original documents, then data entry
- Ponds and flaring
  - Skytruth, Shepardstown, WV – use crowd-sourcing and satellite data
"In a tiny office in Shepherdstown, W.Va., one man is changing the environmental movement."

Washington Post Magazine
August 4, 2013

John Amos, David Manthos, Paul Woods
Variables that Can Be Used for Exposure Assessment

- **WELLS**
  - Location
  - Dates: PERMIT, SPUD, PERF, STIM, PROD
  - Other: PRODUCTION, TOTAL DEPTH, location & size of PONDS, whether flared or not

- **COMPRESSORS**
  - Location on approximately 541 from DEP for UNGD
  - Date in service, number and size of compressor engines, dehydrators, generators

- As of June 30, 2013
  - 6857 wells with SPUD date, have location for all
  - 4381 wells with non-zero PRODUCTION
  - 3036 wells with TOTAL DEPTH
  - 2215 wells with PERF and STIM dates

Challenges with Exposure Assessment

- Development is rapidly varying in time and space; exposure assessment must capture this
- Exposure varies by stage of well development; exposure assessment must capture this
- Information on wells and infrastructure does not necessarily capture what we really care about – toxicant exposures, contextual effects
- Delays in reporting, data entry, data availability
- Identification of the population of UNGDO wells
- Missing information
- Inconsistent information in different sources
- Changes in information over time
Thank you for listening

Questions?

My email:

bschwart@jhsph.edu