ASPO Survivorship, Health Outcomes & Comparative Effectiveness Research Special Interest Group presents:

Connecting the dots: the creation and use of data linkages to study cancer survivorship and health outcomes research

The availability of “big data” in cancer research suggest the potential for new answers to long-standing questions about how to best deliver high quality care to cancer survivors. However, there are several challenges to creating linkages and caveats for their use. This webinar will provide case studies of the creation of new data linkages and discuss how best to leverage existing linkages for cancer control research.

Date: January 8, 2019
Time: 2:00pm-3:00pm ET
Link: https://cwru.zoom.us/j/929650435

Erin Kent, PhD
National Cancer Institute

Stephanie Wheeler, PhD, MPH
University of North Carolina, Chapel Hill

Betsy Shenkman, PhD
University of Florida

Karen Wernli, PhD
Kaiser Permanente

Kate Weaver, PhD MPH
Wake Forest University
NCI Linked Data Resources for Cancer Survivors

Publicly Accessible Data

- Medical Expenditure Panel Survey: Experiences with Cancer Survivorship Supplement (MEPS)
- National Health Interview Survey (NHIS) Cancer Control Supplement (CCS)
- SEER-CAHPS Linked Data Resource
- SEER-Medicare Linked Database
- SEER-MHOS Linked Data Resource

https://healthcaredelivery.cancer.gov/
SEER-MHOS Linked Data Resource
Surveillance, Epidemiology and End Results – Medicare Health Outcomes Survey

Survey includes:
- Health-related quality of life (SF-36, VR-12)
- Activities of daily living
- HEDIS effectiveness of care
- Patient-reported outcomes relevant for older adults with cancer

- Over 140,000 SEER-linked Medicare Advantage (HMO) beneficiaries*
- Over 2 million beneficiaries without cancer
- Designed to be longitudinal with baseline and follow-up surveys, spaced two years apart, proportion with surveys before and after dx
- Over 80 data use agreements and 45+ publications since 2010 launch

*No healthcare claims available in SEER-MHOS of medical care; Part D prescription drug claims under investigation though a feasibility study

https://healthcaredelivery.cancer.gov/seer-mhos/
SEER-CAHPS Linked Data Resource

Surveillance, Epidemiology and End Results – Consumer Assessment of Healthcare Providers and Systems

Cancer Registry Data (SEER)

Healthcare Claims (Medicare)

Patient reported experiences of care surveys (CAHPS)

CAHPS survey includes:

- Doctor Communication
- Getting Needed Care
- Getting Care Quickly
- Care Coordination

- Over 205,000 cancer respondents
- More than 724,000 non-cancer respondents
- Medicare claims allow examination of aspects of healthcare utilization
- Over 10 data use agreements and 8+ publications since 2015 launch
- Rich opportunities for research on patient experiences in cancer care delivery

https://healthcaredelivery.cancer.gov/seer-cahps/

NATIONAL CANCER INSTITUTE
ASPO 2019 Survivorship, Health Outcomes, and Comparative Effectiveness SIG
March 11 Breakfast Session:

**Multiple chronic conditions and care coordination among cancer survivors**

*Individuals are living longer with a history cancer, and many are dealing with other chronic conditions in addition to late/long-term effects from cancer and cancer treatment. The need to manage and coordinate health care services and communicating with a variety of healthcare providers can be challenging. This year’s SIG will focus on research questions related to the coordination of care delivery for cancer patients with multiple chronic conditions.*
ASPO Survivorship, Health Outcomes & Comparative Effectiveness Research Special Interest Group presents:

Connecting the dots: the creation and use of data linkages to study cancer survivorship and health outcomes research

I. Stephanie Wheeler, PhD, MPH, UNC Chapel Hill
   *Getting into the weeds: State-level data linkages for cancer prevention and control research*

II. Betsy Shenkman, PhD, University of Florida
    *OneFlorida Cancer Control Alliance: leveraging linked private and public data for observational and clinical trials*

III. Karen Wernli, PhD, Kaiser Permanente
    *Using Optum claims data in US cancer patients: an example in adolescent and young adults*

Discussant, Q & A: Kate Weaver, PhD MPH, Wake Forest University
Inclusive of Diverse Populations

What Works Better for Whom Under What Circumstances?

Conducted in Real World Settings to Accelerate Translation to Practice and Adoption

The Science of Fostering the Uptake of Evidence-Based Best Practices Into Diverse Health Care Settings

Learning Health System
Stakeholder Engagement: People, Clinicians, Health Systems, Payers, Industry
Real-World Evidence — What Is It and What Can It Tell Us?
DOI: 10.1056/NEJMsb1609216

FDA: Real World Evidence: Why is this happening now?

Real-World Evidence and Real-World Data for Evaluating Drug Safety and Effectiveness
HEALTH CARE SYSTEMS
& AFFILIATED PRACTICES

- University of Florida and UF Health
- Florida State University
  and Regional Campus Practice Partners
- University of Miami and UHealth
- Orlando Health System
- Florida Hospital/AdventHealth
- Tallahassee Memorial HealthCare
- Health Choice Network
- Bond Community Health Center Inc.
- Nicklaus Children’s Hospital
- Community Health IT
- Capital Health Plan

STATE AGENCY & ORGANIZATION COLLABORATIONS

Florida Agency for Health Care Administration
Florida Department of Health
Florida Association of Children’s Hospitals
Encounters of Completed Data

Diagnoses

Procedures

Dispensed Medications
Over 1.7M patients are de-duplicated

<table>
<thead>
<tr>
<th></th>
<th>UFH</th>
<th>ORL</th>
<th>UMI</th>
<th>FLM</th>
<th>TMA</th>
<th>TMC</th>
<th>AVH</th>
<th>NCH</th>
<th>CHP</th>
</tr>
</thead>
<tbody>
<tr>
<td>UFH</td>
<td>X</td>
<td>19,416</td>
<td>13,846</td>
<td>492,255</td>
<td>13,148</td>
<td>28,156</td>
<td>49,861</td>
<td>3150</td>
<td>235</td>
</tr>
<tr>
<td>ORL</td>
<td>19,416</td>
<td>X</td>
<td>3,331</td>
<td>136,315</td>
<td>789</td>
<td>2,526</td>
<td>211,875</td>
<td>1,924</td>
<td>2</td>
</tr>
<tr>
<td>UMI</td>
<td>13,846</td>
<td>3,331</td>
<td>X</td>
<td>205,274</td>
<td>800</td>
<td>2,034</td>
<td>8,124</td>
<td>52,093</td>
<td>3</td>
</tr>
<tr>
<td>FLM</td>
<td>492,255</td>
<td>136,315</td>
<td>205,274</td>
<td>X</td>
<td>43,131</td>
<td>98,291</td>
<td>357,096</td>
<td>238,686</td>
<td>244</td>
</tr>
<tr>
<td>TMA</td>
<td>13,148</td>
<td>789</td>
<td>800</td>
<td>43,131</td>
<td>X</td>
<td>121,697</td>
<td>1,570</td>
<td>424</td>
<td>1,509</td>
</tr>
<tr>
<td>TMC</td>
<td>28,156</td>
<td>2,526</td>
<td>2,034</td>
<td>98,291</td>
<td>121,697</td>
<td>X</td>
<td>4,777</td>
<td>863</td>
<td>2,077</td>
</tr>
<tr>
<td>AVH</td>
<td>49,861</td>
<td>211,875</td>
<td>8,124</td>
<td>357,096</td>
<td>1,570</td>
<td>4,777</td>
<td>X</td>
<td>3,424</td>
<td>9</td>
</tr>
<tr>
<td>NCH</td>
<td>3150</td>
<td>1,924</td>
<td>52,093</td>
<td>238,686</td>
<td>424</td>
<td>863</td>
<td>3,424</td>
<td>X</td>
<td>0</td>
</tr>
<tr>
<td>CHP</td>
<td>235</td>
<td>2</td>
<td>3</td>
<td>244</td>
<td>1,509</td>
<td>2,077</td>
<td>9</td>
<td>0</td>
<td>X</td>
</tr>
</tbody>
</table>
Uses i2b2 software over the Data Trust data to enable researchers to conduct anonymous queries on their own.
• **Aim 1 (Use Characteristics).** In a cohort of patients with an invasive single primary solid tumor, describe the use of common molecular tumor and, in some cases, germline biomarker testing and associated targeted cancer therapies.

• **Aim 2 (Test Results).** In a subcohort of patients identify those who had molecular biomarker testing and for whom a test result was available, and determine whether the selected treatment was in accordance with the test result.

• **Aim 3 (Completeness and Outcomes).** Using the cohort from Aim 1 in sites with linked claims data, assess the completeness of the electronic health record derived data for identifying cancer treatments.
Goal to test interventions in diverse real world settings

- Used Data Trust to Identify Clinic Settings Considering
  - Vaccine Rates
  - Urban/Rural
  - Numbers of Teens, Young Adults
Using OneFlorida Data Trust

Initial Recruitment: UM and UF (Gainesville and Jacksonville)

Expansion to Advent Health
Getting into the weeds: State-level data linkages for cancer prevention and control research

Stephanie Wheeler, PhD MPH
Lineberger Comprehensive Cancer Center
Health Policy & Management
Gillings School of Global Public Health
University of North Carolina at Chapel Hill
Objectives

- Describe statewide data linkages being used for cancer prevention and control research
- Highlight how statewide linked cancer data can be used to identify population health problems and target potential solutions across the continuum
- Summarize with lessons learned/best practices for linking and leveraging state cancer data
What do I mean by statewide cancer data linkages?

• Direct or probabilistic linkages among two or more secondary datasets that contain statewide cancer-related data:
  – Cancer surveillance (registry) data
  – Health insurance administrative claims and billing data
  – Healthcare resource, facility, and workforce data
  – Hospital discharge data
  – ED utilization data
  – Immunization records
  – DMV data
  – Sociodemographic and economic data
  – Bankruptcy filings
Developing Real World Linked Cancer Data Resources

Integrated, Inter-disciplinary team science
- Clinical domain experts
- Population/public health scientists
- Computer scientists/programmers
- Statisticians
- Database analysts

Steering Committee
Faculty and Collaborators
Research Staff
Software Developers

CIPHR Team

Data
- Cancer Registry Data
- Private and Public Payer Data
- Provider Information
- Geo-Spatial Resources

Methods
- Patient Matching
- Cohort Creation
- Clinical Context Identification
- Longitudinal Data Interpretation

Systems
- Secure Computing Systems
- Web Systems
- Regulatory Security Compliance

UNC Lingleberger
**UNC Lineberger Cancer Information & Population Health Resource (CIPHR)**

**Unique linkages:**
Cancer registry, multi-payer claims data (100% Medicare, 100% Medicaid, 70% private), SSI death index, other contextual data

**Health Care Claims:**
>6m persons since 2003
55% of NC population

**NC Cancer Registry:**
100% since 2003-2015
>650,000 cases

**Cancer-cases claims:**
85% of NC cancers
>552,000

**Key collaborators**
Chris Baggett
Laura Green
May Kuo
Public Health Faculty

**Medicine Faculty**

**Shared resources**
4 Systems developers
6 Analysts
1 program coordinator

**Key pubs (>60)**
Meyer et al, NCMJ, 2014
Wheeler et al, H&P, 2014
Wheeler et al, Medical Care, 2013
Unique linkages:
Multipayer claims data (Medicaid, private insurers), other contextual data

Health Care Claims:
From 2007 for Medicaid; 2010 for private

Key collaborators
John McConnell
Stephanie Renfro
Bonnie Lind
Public Health Faculty
Medicine Faculty

Shared resources
3 Health economists
5 Statisticians
3 Research assistants
1 program coordinator

Key pubs (>52)
McConnell et al, Health Affairs, 2017
Davis et al, J of Rural Health, 2016
Charlesworth et al, JAMA IM, 2016
Examples of Contextual Data to Augment State-level Cancer Registry-Linked-Claims

- Area Resource Files (ARF)
- Census/American Community Survey
- State Medical Facilities Plan Data
- State Pharmacy Association Data
- State Oncology Association Data
- RWJ County Health Rankings
- Public Health and Primary Care Networks
- National Association of County and City Health Officials (NAACHO)
Geolocating Endoscopy Facilities in NC
Geolocating Community Pharmacies in NC
Geolocating American Cancer Society Primary Care Managers in NC
How Can State Data Linkages Be Useful?

• Identify geographic, socio-demographic and temporal trends in cancer risk factors (e.g., HPV vax), screening, incidence, mortality, care quality (e.g., surgery, chemo, radiation therapy), costs
• Identify and locate healthcare resources
• Identify social determinants of health outcomes
• Enable multilevel multivariable statistical and simulation modeling
• Target specific regions or sub-populations for interventions/implementation
Identifying geographic and socio-demographic correlates of colorectal cancer screening

Geographic and population-level disparities in colorectal cancer testing: A multilevel analysis of Medicaid and commercial claims data

Melinda M. Davis a,b, Stephanie Renfro a, Robyn Pham a,b, Kristen Hassmiller Lich c,d, Jackilen Shannon a, Gloria D. Coronado e, Stephanie B. Wheeler d,a,h

Regional variation in colorectal cancer testing and geographic availability of care in a publicly insured population

Stephanie B. Wheeler a, b, c, d, Tzy-Mey Kuo b, Ravi K. Goyal b, Anne-Marie Meyer b, Kristen Hassmiller Lich a, Emily M. Gillan a, Seth Tyro e, Carmen L. Lewis b, c, Trisha M. Crutchfield d
Identifying geographic and socio-demographic correlates of cancer treatment access & quality

Original Investigation

November 2017

Association of Delays in Surgery for Melanoma With Insurance Type

Adewole S. Adamson, MD, MPP; Lei Zhou, MSPH; Christopher D. Baggett, PhD; et al.

Cancer

Original Article

Influence of provider factors and race on uptake of breast cancer gene expression profiling

Katherine E. Reeder-Hayes MD, MBA; Stephanie B. Wheeler PhD, Christopher D. Baggett PhD; Xi Zhou MS, Ke Meng PhD, Megan C. Roberts PhD, Lisa A. Carey MD, Anne-Marie Meyer PhD

Medical Care. 56(5):A30–A35, MAY 2018
DOI: 10.1097/MLR.0000000000000906, PMID: 29578593

Care Coordination and Multispecialty Teams in the Care of Colorectal Cancer Patients

Justin G. Tredgon; Yunkyoung Chang; Sarah Shai; Peter J. Mucha; Tzy-Mey Kuo; Anne M. Meyer; Karyn B. Stitzenberg
Identifying cancer care costs and cost-effectiveness of treatments and interventions

Medical costs of treating breast cancer among younger Medicaid beneficiaries by stage at diagnosis

Justin G. Trogdon1, Donatus U. Ekwueme2, Diana Poehler3, Cheryll C. Thomas2, Katherine Reeder-Hayes1, and Benjamin T. Allaire3

Cost Analysis of Robot-Assisted Laparoscopic Versus Hand-Assisted Laparoscopic Partial Nephrectomy

James E. Ferguson3, Rani K. Goyal, Matthew C. Raynor, Matthew E. Nielsen, Raj S. Pruthi, Paul M. Brown, and Eric M. Wallet

Breast cancer treatment costs in younger, privately insured women

Benjamin T. Allaire1, Donatus U. Ekwueme2, Diana Poehler1, Cheryll C. Thomas2, Gery P. Guy Jr.2, Sujha Subramanian1, and Justin G. Trogdon3

Cost-Effectiveness Analysis of Four Simulated Colorectal Cancer Screening Interventions, North Carolina

Kristen Hassmiller Lich, PhD1; David A. Cornejo2; Maria E. Mayorga, PhD5; Michael Pignone, MD, MPH1,4,5; Florence K.L. Tangka, PhD1; Lisa C. Richardson, MD, MPH1; Tzy-Mey Kuo, PhD, MPH1; Anne-Marie Meyer, PhD4,5; Ingrid J. Hall, PhD, MPH1; Judith Lee Smith, PhD1; Todd A. Durham, MS1; Steven A. Chail, MS3; Trisha M. Crotchfield, MHA, MSIS4,6; Stephanie B. Wheeler, PhD, MPH1,4,5

First published: 19 August 2015 | https://doi.org/10.1002/cncr.29596 | Cited by: 4
Enabling multilevel multivariable statistical and simulation modeling to project outcomes

Population

Screening Patterns

Cancer Progression

Cancer Outcomes

Policy Effects

NC-CRC Simulation Model
Geo-spatially explicit, population-based, individual-level simulation model of natural history of CRC and screening behaviors

Population

Census Data
American Community Survey

Claims Data
Medicare, Medicaid, Private Insurance

RTI Model
Natural history of adenomas and cancer

Cancer Registry
Population-based CRC incidence

Literature Review,
American Community Survey, BRFSS

Synthetic Population
Realistic population of individuals

Statistical Models
Regression models predicting modality and compliance

Parameter Estimates
Calibration

Predicted Probabilities

Parameter Estimates

Structural assumptions and parameter values

Population inputs
Targeting specific regions or sub-populations for interventions/implementation

Cancer

Comparative effectiveness of mailed reminders with and without fecal immunochemical tests for Medicaid beneficiaries at a large county health department: A randomized controlled trial

Excluded = 418
- Bad Address = 224
- Previous COLO = 92
- Previous FOBT = 10
- Previous Other = 10
- Opted Out = 82

Reminder + FIT
- N = 1071
  Included 716
  Returned FIT 151 (21%)
    Positive 11
    Negative 136

Reminder ONLY
- N = 1073
  Included 655
  Requested FIT 147
  Returned FIT 85 (13%)
    Negative 79
    Positive 6

Difference 8% (4%, 12%; p<0.01)
Lessons Learned

• Dedicate resources to build data computing infrastructure, expertise & capacity
• Plan (& pay) for regular data updates
• Partner early and be a good partner! (eg, ROI)
• Explore probabilistic data linkages (SSN and name/address not always necessity)
• Develop multidisciplinary teams
• Consider unusual linkages (e.g., retail, environmental, financial, education, transportation and labor market data)
Thank you!
For more info, check out:
https://ciphr.unc.edu
Contact me at:
stephanie_wheeler@unc.edu
@StephWheelerUNC
Using Optum claims data in US cancer patients: an example in adolescent and young adults

Karen J. Wernli, PhD
Associate Investigator
January 8, 2019
Financial Disclosures

• Nothing to disclose
Challenge that I faced
Current R21 specific aims

To assess variation in end of life among adolescent and young adult cancer patients, from 2001 – 2016 by time and geography, evaluating:

- Emergency department visits
- Hospitalizations
- Intensive Care Unit stays
- Surgery
- Receipt of chemotherapy
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 30 days of death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2 ED visits</td>
<td>16.3%</td>
<td>15.3%</td>
<td>15.6%</td>
<td>16.3%</td>
<td>22.1%</td>
<td>0.01</td>
</tr>
<tr>
<td>Hospitalization</td>
<td>78.4%</td>
<td>-</td>
<td>77.3%</td>
<td>79.1%</td>
<td>80.7%</td>
<td>0.24</td>
</tr>
<tr>
<td>ICU stay</td>
<td>40.1%</td>
<td>35.8%</td>
<td>40.2%</td>
<td>42.7%</td>
<td>42.9%</td>
<td>0.004</td>
</tr>
<tr>
<td>Surgery</td>
<td>4.5%</td>
<td>4.4%</td>
<td>4.1%</td>
<td>5.1%</td>
<td>4.5%</td>
<td>0.63</td>
</tr>
<tr>
<td>Within 14 days of death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>12.0%</td>
<td>12.5%</td>
<td>11.5%</td>
<td>12.0%</td>
<td>12.3%</td>
<td>0.90</td>
</tr>
<tr>
<td>Measure</td>
<td>Overall</td>
<td>Northeast</td>
<td>Midwest</td>
<td>South</td>
<td>West</td>
<td>Chi-square p-value</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
<td>------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within 30 days of death</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2 ED visits</td>
<td>16.3%</td>
<td>16.7%</td>
<td>16.6%</td>
<td>15.9%</td>
<td>16.9%</td>
<td>0.92</td>
</tr>
<tr>
<td>Hospitalization</td>
<td>78.4%</td>
<td>83.4%</td>
<td>78.2%</td>
<td>79.1%</td>
<td>74.3%</td>
<td>0.008</td>
</tr>
<tr>
<td>ICU stay</td>
<td>40.1%</td>
<td>41.9%</td>
<td>36.2%</td>
<td>42.6%</td>
<td>38.3%</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td>Surgery</td>
<td>4.5%</td>
<td>5.8%</td>
<td>4.0%</td>
<td>4.6%</td>
<td>4.5%</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within 14 days of death</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>12.0%</td>
<td>13.0%</td>
<td>11.7%</td>
<td>12.2%</td>
<td>11.5%</td>
<td>0.87</td>
</tr>
</tbody>
</table>
What are these claims data good for?

- Healthcare data for commercially insured <65 years in all states and over time
- Patterns of healthcare utilization: time, geography
- Evaluation of change in care: chemotherapy, hospital
- Short-term outcomes
- Evaluating mortality
Caveats

- Lacking cancer diagnosis date, stage and tumor type
- Longitudinal analyses
- Limited confounders
Process

• In grant preparation, began speaking to Optum in support of purchasing claims data
• Optum provided a letter of support for application
• Work with an Optum programmer to build a cohort dataset
• Optum provides training of data to research team
• Optum available for additional data support as needed
Thank you

Karen.J.Wernli@kp.org
206-287-2934
@WernliKarenJ