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



https://healthcaredelivery.cancer.gov/

SEER-MHOS Linked Data Resource

Surveillance, Epidemiology and End Results – Medicare Health Outcomes Survey

Cancer
Registr
y Data
(SEER)
Patient
health
outcome
s survey
(MHOS)

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)
Patient
Healthca reported
experien
ces 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/

https://aspo.org/annual-meeting/



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

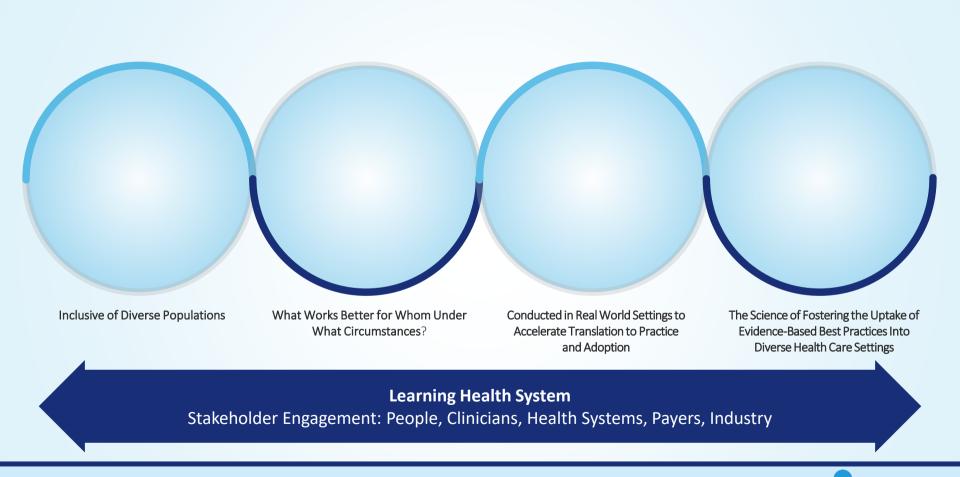


III. Karen Wernli, PhD, Kaiser Permanente
Using Optum claims data in US cancer patients: an example in adolescent and young adults



Discussant, Q & A: Kate Weaver, PhD MPH, Wake Forest University





Real-World Evidence — What Is It and What Can It Tell Us?

N Engl J Med 2016; 375:2293-2297

DOI: 10.1056/NEJMsb1609216

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



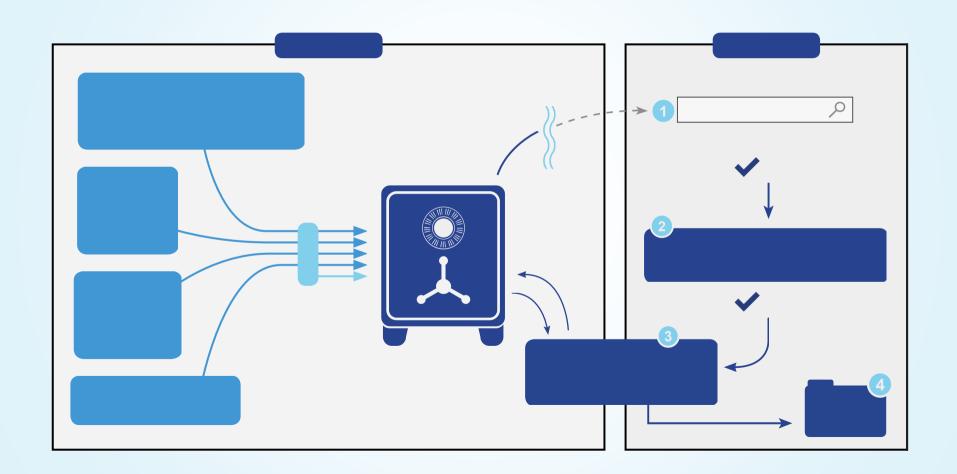
Real-World Evidence and Real-World Data for Evaluating Drug Safety and Effectiveness

JAMA. 2018;320(9):867-868. doi:10.1001/jama.2018.10136









FUNDAMENTAL DATA

Demographic

- PATID
- BIRTH DATE
- BIRTH TIME
- SFX
- HISPANIC
- RACE
- BIOBANK FLAG

HEALTH CARE DELIVERY DATA

Enrollment

- PATID
- ENR START DATE
- ENR END DATE
- CHART
- ENR BASIS Dispensing

- DISPENSINGID
- PATID PRESCRIBINGID (optioal)
- DISPENSE DATE
- NDC
- DISPENSE SUP DISPENSE AMTn

Death

- PATID
- DEATH DATE
- DEATH DATE IMPUTE
- DEATH SOURCE
- DEATH MATCH CONFIDENCE

Death Conditio

- PATID
- DEATH CAUSE
- DEATH CAUSE CODE
- DEATH_CAUSE_TYPE **DEATH CAUSE SOURCE**
- **DEATH CAUSE CONFIDENCE**

MULTIPLE CONTEXT DATA

Vital

- VITALID
- ΡΔΤΙΠ
- ENCOUNMERID (optioal)
- MEASURE DATE MEASURE TIME
- VITAL SOURCE
- HT
- WT
- DIASTOLIC
- SYSTOLIC ORIGINAL BMI
- BP POSITION
- SMOKING
- TOBACCO
- TOBACCO TYPE

Conditio

- CONDITIONID
- PATID
- ENCOUNMEREDID (optioal)
- REPORT DATE
- RESOLVE DATE ONSET DATE
- CONDITION STATUS
- CONDITION
- CONDITION TYPE
- CONDITION SOURCE

Pro CM

- PRO CM ID
- PATID
- ENCOUNMEREDID (optioal) PRO ITEM
- PRO LOINC
- PRO DATE
- PRO TIME
- PRO RESPONSE PRO METHOD
- PRO MODE PRO CAT

DIRECT ENCOUNTER DATA

Encounter

- ENCOUNTERID
- PATID
- ADMIT DATE
- ADMIT TIME DISCHARGE DATE
- DISCHARGE TIME PROVIDERID
- FACILITY LOCATION
- ENC TYPE FACILITYID
- DISCHARGE DISPOSITION
- DISCHARGE STATUS DRG
- DRG TYPE
- ADMITTING SOURCE Diagnosis
- DIAGNOSISID
- PATID
- ENCOUNTERID ENC TYPE
- ADMIT DATE
- PROVIDERID
- PX DATE PX
- PX TYPE
- PX SOURCE **Procedures**

PROCEDURESID

- PATID
- ENCOUNTERID
- ENC TYPE ADMIT DATE
- PROVIDERID
- PX DATE PX
- PX TYPE PX SOURCE

- Lab Result
- LAB RESULT CM ID PATID
- FNCOUNTERID
- LAB NAME
- SPECIMEN SOURCE LAB LOINC
- PRIORITY RESULT LOC LAB PX
- LAB PX TYPE LAB ORDER DATE
- SPECIMEN DATE SPECIMEN TIME
- RESULT DATE RESULT TIME
- RESULT QUAL
- RESULT NUM RESULT MODIFIER
- RESULT UNIT NORM RANGE LOW
- NORM MODIFIER LOW NORM RANGE HIGH
- NORM MODIFIER HIGH ABN IND
- Prescribing PRESCRIBINGID
- PATID ENCOUNTERID
- RX PROVIDERID
- RX ORER DATE RX ORDER TIME
- RX START DATE RX END DATE RX QUANTITY
- RX REFILLS RX DAYS SUPPLY
- RX FREQUENCY RX BASIS RXNORM CUI

PCORnet TRIAL DATA

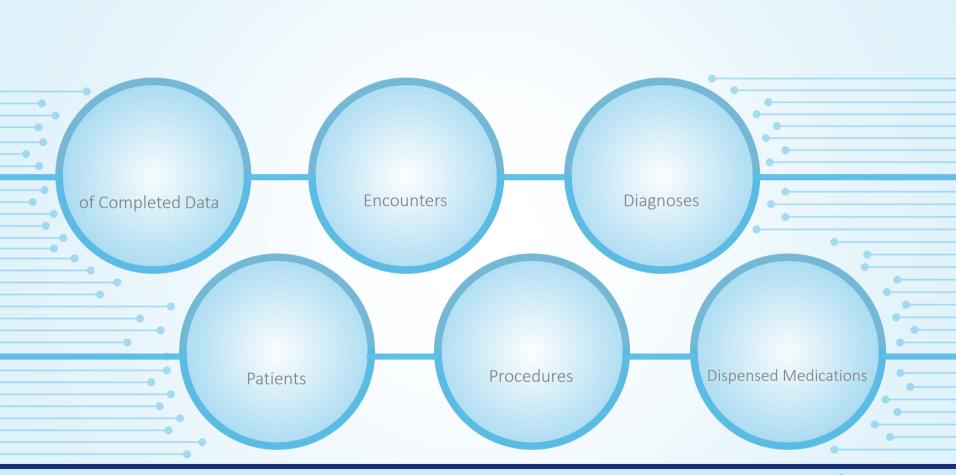
PCORnet Trial

- PATID
- TRAILID
- PARTICIPANTID
- TRIAL SITEID
- TRIAL ENROLL DATE
- TRIAL END DATE
- TRIAL WITHDRAW DATE TRIAL INVITE CODE

PROCESS-RELATED DATA

Harvest

- NETWORKID NETWORK NAME
- DATAMARTID
- DATAMART NAME DATAMART PLATFORM CDM VERSION
- DATAMART CLAIMS
- DATAMART EHR BIRTH DATE MGMT
- ENR START DATE MGMT
- ENR END DATE MGMT
- ADMIT DATE MGMT DISCHARGE DATE MGMT
- PX DATE MGMT RX ORDER DATE MGMT
- RX START DATE MGMT
- RX END DATE MGMT DISPENSE DATE MGMT
- LAB ORDER DATE MGMT SPCIMEN DATE MGMT
- RESULT DATE MGMT MEASURE DATE MGMT
- ONSET DATE MGMT REPORT DATE MGMT
- RESOLVE DATE MGMT PRO DATE MGMT
- REFRESH DEMOGRAPHIC DATE REFRESH ENROLLMENT DATE
- REFRESH ENCOUNTER DATE REFRESH DIAGNOSIS DATE
- REFRESH PROCEDURES DATE REFRESH VITAL DATE
- REFRESH DISPENSING DATE
- REFRESH LAB RESULT CM DATE REFRESH CONDITION DATE
- REFRESH PRO CM DATE REFRESH PRESCRIBING DATE
- REFRESH PCORNET TRIAL DATE REFRESH DEATH DATE
- REFRESH CAUSE DEATH DATE





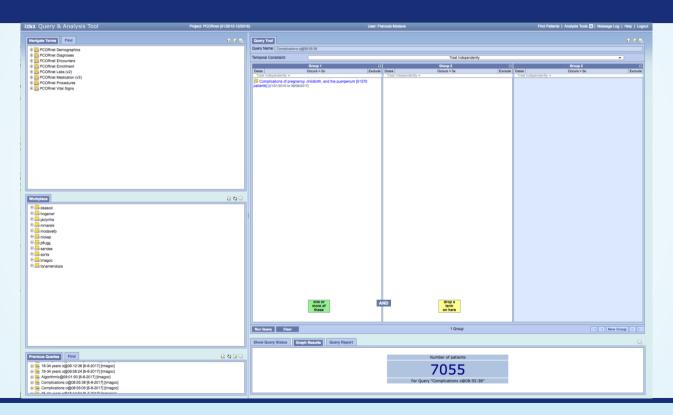
Over 1.7M patients are de-duplicated

UFH	X	19,416	13,846	492,255	13,148	28,156	49,861	3150	235
ORL	19,416	X	3,331	136,315	789	2,526	211,875	1,924	2
UMI	13,846	3,331	Χ	205,274	800	2,034	8,124	52,093	3
FLM	492,255	136,315	205,274	X	43,131	98,291	357,096	238,686	244
TMA	13,148	789	800	43,131	Χ	121,697	1,570	424	1,509
TMC	28,156	2,526	2,034	98,291	121,697	X	4,777	863	2,077
AVH	49,861	211,875	8,124	357,096	1,570	4,777	Χ	3,424	9
NCH	3150	1,924	52,093	238,686	424	863	3,424	X	0
CHP	235	2	3	244	1,509	2,077	9	0	X



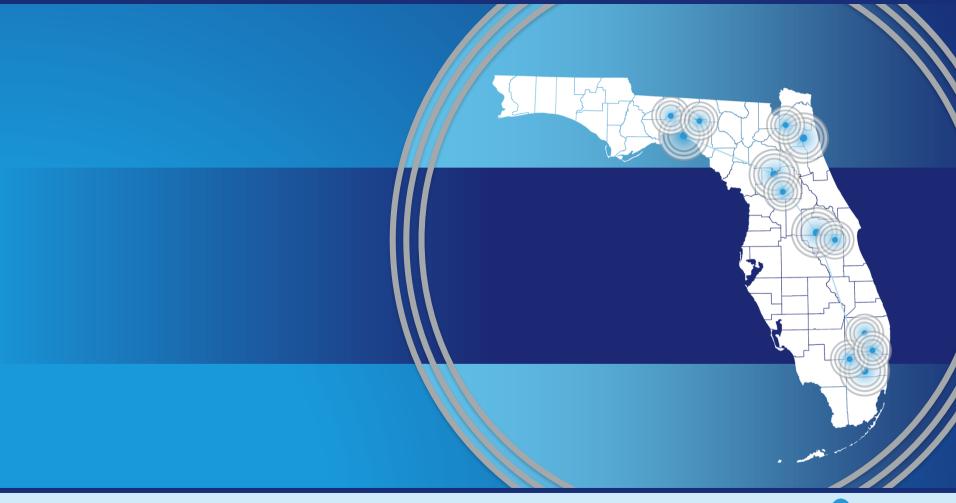
. . .

Uses i2b2 software over the Data Trust data to enable researchers to conduct anonymous queries on their own.

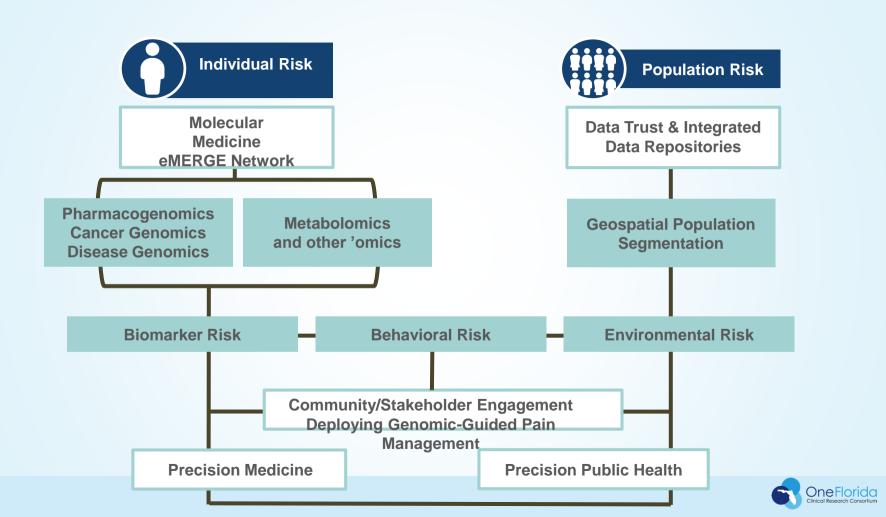




• • • •







PCORI Rapid Cycle Research: Patterns of Use of Molecular Biomarkers and Targeted Cancer Therapies

- 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

NCI: Improving the Uptake of HPV Vaccine PI: Stephanie Staras, PhD

- 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?

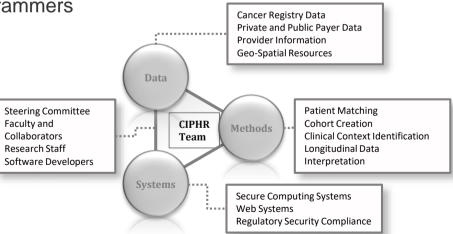
- Direct or probabilistic linkages among two or more secondary datasets that contain statewide cancer-related data:
 - Cancer surveillance (registry) data
 - Health insurance administrative claims and billing data
 - Healthcare resource, facility, and workforce data
 - Hospital discharge data
 - ED utilization 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





UNC Lineberger Cancer Information & Population Health Resource (CIPHR)

Unique linkages:

contextual data

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

Health Care Claims:

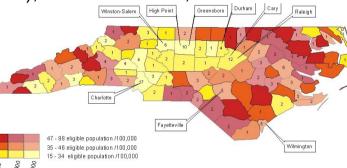
>6m persons since 2003 55% of NC population

NC Cancer Registry:

100% since 2003-2015 >650,000 cases

Cancer-cases claims:

85% of NC cancers >552,000



* Number showing in each county are the number of endoscopy facilities

Key collaborators

Chris Baggett
Laura Green
May Kuo
Public Health

Faculty

Medicine Faculty

Shared resources

4 Systems developers 6 Analysts 1 program

Key pubs (>60)

coordinator

Meyer et al, NCMJ, 2014

Wheeler et al, H&P, 2014

Wheeler et al, Medical Care, 2013





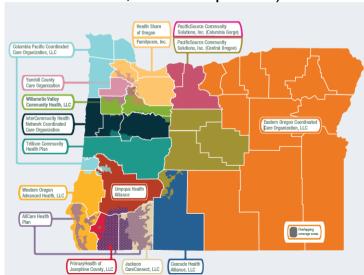
Oregon Health & Sciences University Center for Health Systems Effectiveness

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

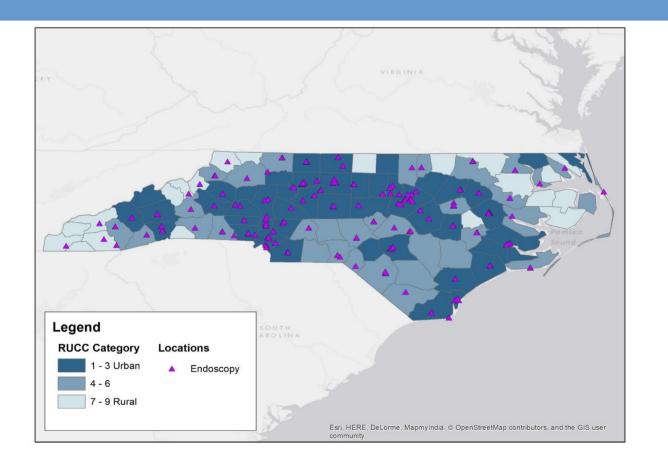


Examples of Contextual Data to Augment State-level Cancer Registry-Linked-Claims

- 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)

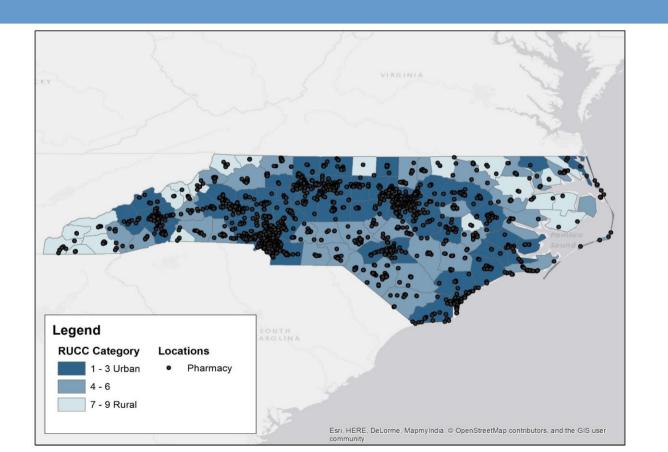


Geolocating Endoscopy Facilities in NC





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



Preventive Medicine

Available online 13 May 2017

In Press, Accepted Manuscript - Note to users



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^{c, ™}, Robyn Pham^{b, ™}, Kristen Hassmiller Lich^{d, ™}, Jackilen Shannon^{e, ™}, Gloria D. Coronado^{f, ™}, Stephanie B. Wheeler^{d, g, h, ™}



Health & Place

Volume 29, September 2014, Pages 114-123



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



Identifying geographic and socio-demographic correlates of cancer treatment access & quality



Urologic Oncology: Seminars and Original Investigations

Volume 36, Issue 6, June 2018, Pages 308.e1-308.e9

Original article

The relationship of travel distance with cystectomy access and outcomes &

Angela B. Smith M.D., M.S. a, b ≥ M. Anne-Marie Meyer Ph.D. c, Ke Meng Ph.D. c, Matthew E. Nielsen M.D., M.S. a, b, d, Raj Pruthi M.D. a, b, Eric Wallen M.D. a, b, Michael Woods M.D. a, b, Hung-Jui Tan M.D., M.S. a, b

Gynecologic Oncology 152 (2019) 112-118



Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Original Investigation

November 2017

Association of Delays in Surgery for **Melanoma With Insurance Type**

Adewole S. Adamson, MD, MPP^{1,2,3}; Lei Zhou, MSPH³; Christopher D. Baggett, PhD^{3,4}; et al

> Author Affiliations | Article Information

JAMA Dermatol. 2017;153(11):1106-1113. doi:10.1001/jamadermatol.2017.3338



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

First published: 16 January 2018 | https://doi.org/10.1002/cncr.31222

Evaluating the urban-rural paradox: The complicated relationship between distance and the receipt of guideline-concordant care among cervical cancer patients

Lisa P. Spees a,*, Stephanie B. Wheeler b,c, Mahesh Varia d, Morris Weinberger b, Christopher D. Baggett c,e, Xi Zhou ^c, Victoria M. Petermann ^f, Wendy R. Brewster ^{c,e,g}

- ^a The Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, United States of America
- b Department of Health Policy and Management, University of North Carolina at Chapel Hill, United States of America ^c Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, United States of America d Department of Radiation Oncology, University of North Carolina at Chapel Hill, United States of America
- ^e Department of Epidemiology, University of North Carolina at Chapel Hill, United States of America

School of Nursing, University of North Carolina at Chapel Hill, United States of America 8 Department of Obstetrics & Gynecology, University of North Carolina at Chapel Hill, United States of America Medical Care. 56(5):430-435, MAY 2018

DOI: 10.1097/MLR.0000000000000906. PMID: 29578953

Publication Date: 2018/05/01





FREE



Identifying cancer care costs and costeffectiveness of treatments and interventions

Published in final edited form as:

Breast Cancer Res Treat. 2017 November; 166(1): 207-215. doi:10.1007/s10549-017-4386-2.

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

Published in final edited form as:

Published in final edited form as: Breast Cancer Res Treat. 2017 July; 164(2): 429–436. doi:10.1007/s10549-017-4249-x.

Journal of Endourology, Vol. 26, No. 8 | Laparoscopy and Robotic Surgery

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

James E. FergusonIII 🖂 Ravi K. Goyal, Mathew C. Raynor, Matthew E. Nielsen, Raj S. Pruthi, Paul M. Brown, and Eric M. Wallen

Published Online: 7 Aug 2012 https://doi.org/10.1089/end.2011.0568

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



ORIGINAL RESEARCH

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

Kristen Hassmiller Lich, PhD¹; David A. Cornejo²; Maria E. Mayorga, PhD²;
Michael Pignone, MD, MPH^{3,4,5,6}; Florence K.L. Tangka, PhD⁷;
Lisa C. Richardson, MD, MPH⁷; Tzy-Mey Kuo, PhD, MPH²; Anne-Marie Meyer, PhD^{3,8};
Ingrid J. Hall, PhD, MPH⁷; Judith Lee Smith, PhD⁷; Todd A. Durham, MS¹;
Steven A. Chall, MS⁹; Trisha M. Crutchfield, MHA, MSIS^{4,6};

Stephanie B Wheeler PhD MPH13.4

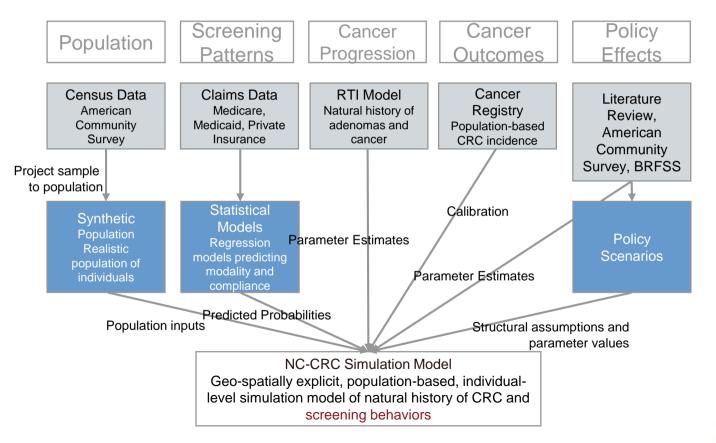


Original Article 🙃 Free Access

Association between medical home enrollment and health care utilization and costs among breast cancer patients in a state Medicaid program

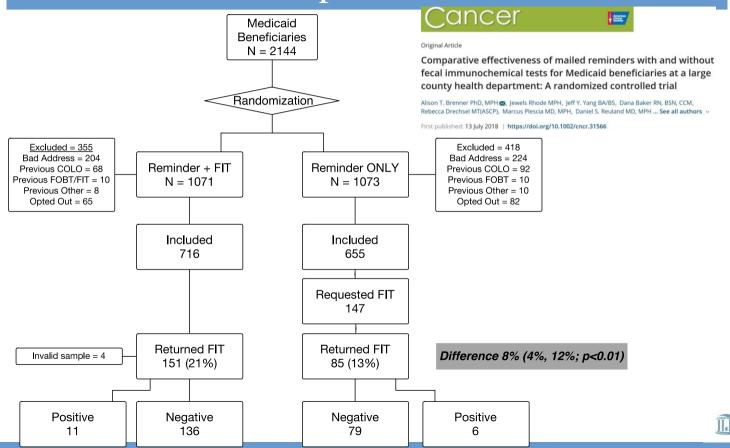
Racquel E. Kohler PhD, Ravi K. Goyal MS, Kristen Hassmiller Lich PhD, MHSA, Marisa Elena Domino PhD, Stephanie B. Wheeler PhD, MPH 🕿

Enabling multilevel multivariable statistical and simulation modeling to project outcomes





Targeting specific regions or sub-populations for interventions/implementation



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



University Cancer Research Fund

















Using Optum claims data in US cancer patients: an example in adolescent and young adults

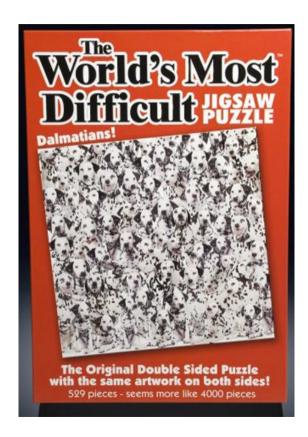
Karen J. Wernli, PhD Associate Investigator January 8, 2019



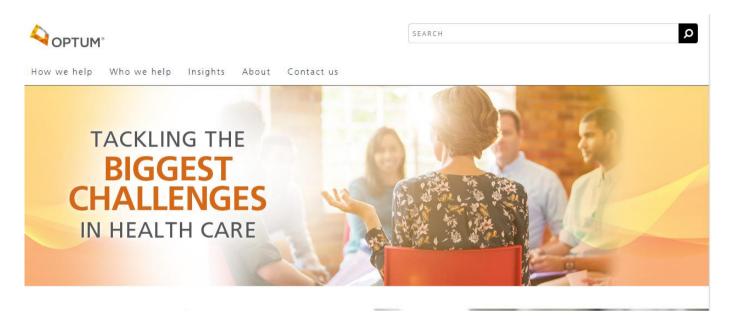
Financial Disclosures

Nothing to disclose

Challenge that I faced



Optum



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 2. Healthcare services at end of life by time period									
Measure	Overa II	2001 2004	2005 2008	2009 2012	2013 2016	Chi-square p-value			
Within 30 days of death									
≥2 ED visits	16.3%	15.3%	15.6%	16.3%	22.1%	0.01			
Hospitalization	78.4%	_	77.3%	79.1%	80.7%	0.24			
ICU stay	40.1%	35.8%	40.2%	42.7%	42.9%	0.004			
Surgery	4.5%	4.4%	4.1%	5.1%	4.5%	0.63			
Within 14 days of death									
Chemotherapy	12.0 %	12.5%	11.5%	12.0%	12.3%	0.90			

TABLE 3. Healthcare services at end of life overall and by region and time									
Measure	Overal I	North- east	Mid- west	South	West	Chi-square p-value			
Within 30 days of death									
≥2 ED visits	16.3%	16.7%	16.6%	15.9%	16.9%	0.92			
Hospitalization	78.4%	83.4%	78.2%	79.1%	74.3%	0.008			
ICU stay	40.1%	41.9%	36.2%	42.6%	38.3%	<0.003			
Surgery	4.5%	5.8%	4.0%	4.6%	4.5%	0.50			
Within 14 days of death									
Chemotherapy	12.0%	13.0%	11.7%	12.2%	11.5%	0.87			

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



