Secondary Uses of Data for Comparative Effectiveness Research

Paul Wallace MD
Director, Center for Comparative Effectiveness Research
The Lewin Group

Paul.Wallace@lewin.com
Disclosure/Perspectives

• Training: Internist/Hematologist/Cancer Research
• Practice: Hematology and Medical Oncology; Population Medicine; EBM/CER/PCOR

Kaiser Permanente (past)
• Heme/Onc Practice
• EBM (Guidelines & Population Management)
• HIT (EMR, Pop’n Mgmt, Oncology Systems)

Lewin/Optum (current)
• Center for CER
  • Inventory of CER and PCOR (ASPE/HHS; AHRQ & NIH)
  • Multi-Payer Claims Database (HHS/ASPE)
• Autism Registry (NIMH)
• Geographic Variation (IOM)
Observations on the data and information infrastructure for cancer...

• This is a very complex area...
  – There is increasing certainty that there are workable and even elegant technical solutions
  – There is minimal agreement about how to scale solutions to a national level and/or sustain them over time

• Hypothesis: The current evolving national effort around comparative effectiveness research is a rich laboratory for exploring and refining effective and durable scaling and sustainability of the informatics support for cancer care and learning
Overview

• Comparative Effectiveness Research (CER) and Patient Centered Outcomes Research (PCOR)
• Secondary use of data
  – Examples
• Sustainability
  – Key issues
  – Conceptual framework
CER

Comparative effectiveness research is the conduct and synthesis of research comparing the benefits and harms of different interventions and strategies to prevent, diagnose, treat and monitor health conditions in “real world” settings...

Federal Coordinating Council for Comparative Effectiveness Research 2009
Medicare Modernization Act 2003; Section 1013 ($15-30M)
“Comparative effectiveness research is the conduct and synthesis of research comparing the benefits and harms of different interventions and strategies to prevent, diagnose, treat and monitor health conditions in “real world” settings”

ARRA 2009 ($1.1B)

AHRQ Effective Health Care Program 2003-Present

ACA 2010

pcori
The **Patient-Centered Outcomes Research Institute (PCORI)** is an independent, non-profit health research organization. Its mission is to fund research that offers patients and caregivers the information they need to make important healthcare decisions.

**PCORI’s Path from Priorities to Research Patients Can Use**
Patient Centered Outcomes Research (PCOR) helps people make informed health care decisions and allows their voice to be heard in assessing the value of health care options.

This research answers patient-focused questions:

- “Given my personal characteristics, conditions and preferences, what should I expect will happen to me?”
- “What are my options and what are the benefits and harms of those options?”
- “What can I do to improve the outcomes that are most important to me?”
- “How can the health care system improve my chances of achieving the outcomes I prefer?”
# PCORI Draft Research Agenda

<table>
<thead>
<tr>
<th>Priority</th>
<th>% Funding Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparative Assessment of Options for Prevention, Diagnosis, and Treatment</td>
<td>Approximately 40%</td>
</tr>
<tr>
<td>Improving Healthcare Systems</td>
<td>Approximately 20%</td>
</tr>
<tr>
<td>Communication and Dissemination Research</td>
<td>Approximately 10%</td>
</tr>
<tr>
<td>Addressing Disparities</td>
<td>Approximately 10%</td>
</tr>
<tr>
<td>Accelerating PCOR and Methodological Research</td>
<td>Approximately 20%</td>
</tr>
</tbody>
</table>
The Evolving Evidence Perspective...

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Methods</th>
<th>Data Source/Organization</th>
<th>Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert Opinion</td>
<td>• Case Reports</td>
<td>• Charts</td>
<td>Effect on Patients (N of 1)</td>
</tr>
<tr>
<td></td>
<td>• Case Series</td>
<td>• Experience</td>
<td></td>
</tr>
</tbody>
</table>

The table above lists different types of studies, their methods, data sources, and their perspectives. The expert opinion study type includes methods such as case reports and case series, with data sources like charts and experience, focusing on the effect on patients (N of 1).
## The Evolving Evidence Perspective...

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Methods</th>
<th>Data Source/Organization</th>
<th>Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert Opinion</td>
<td>• Case Reports</td>
<td>• Charts</td>
<td>Effect on Patients (N of 1)</td>
</tr>
<tr>
<td></td>
<td>• Case Series</td>
<td>• Experience</td>
<td></td>
</tr>
<tr>
<td>Evidence Based Medicine</td>
<td>• RCTs</td>
<td>• Trial Data &amp; Databases</td>
<td>Population Efficacy (N of Many)</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• Meta-analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• (Observation)</td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
<tr>
<td>Study Type</td>
<td>Methods</td>
<td>Data Source/Organization</td>
<td>Perspective</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Expert Opinion</td>
<td>• Case Reports</td>
<td>• Charts</td>
<td>Effect on Patients (N of 1)</td>
</tr>
<tr>
<td></td>
<td>• Case Series</td>
<td>• Experience</td>
<td></td>
</tr>
<tr>
<td>Evidence Based Medicine</td>
<td>• RCTs</td>
<td>• <strong>Trial Data &amp; Databases</strong></td>
<td>Population Efficacy (N of Many)</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• <strong>Meta-analysis</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• (Observation)</td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
<tr>
<td>Comparative Effectiveness Research</td>
<td>• RCTs</td>
<td>• Trial Data &amp; Databases</td>
<td>Population Effectiveness (N of Many)</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• Meta-analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Observation</td>
<td>• Large Pop’n Databases</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
</tbody>
</table>
# The Evolving Evidence Perspective...

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Methods</th>
<th>Data Source/Organization</th>
<th>Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert Opinion</td>
<td>• Case Reports</td>
<td>• Charts</td>
<td>Effect on Patients (N of 1)</td>
</tr>
<tr>
<td></td>
<td>• Case Series</td>
<td>• Experience</td>
<td></td>
</tr>
<tr>
<td>Evidence Based Medicine</td>
<td>• RCTs</td>
<td>• <strong>Trial Data &amp; Databases</strong></td>
<td>Population Efficacy (N of Many)</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• <strong>Meta-analysis</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• (Observation)</td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
<tr>
<td>Comparative Effectiveness Research</td>
<td>• RCTs</td>
<td>• Trial Data &amp; Databases</td>
<td>Population Effectiveness (N of Many)</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• <strong>Meta-analysis</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Observation</td>
<td>• <strong>Large Pop’n Databases</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
<tr>
<td>Patient Centered Outcomes Research</td>
<td>• RCTs</td>
<td>• Trial Data &amp; Databases</td>
<td>Patient Effectiveness Many (N of 1)s</td>
</tr>
<tr>
<td></td>
<td>• Systematic Reviews</td>
<td>• Meta-analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <strong>Systematic Observation</strong></td>
<td>• Large Pop’n Databases</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reports &amp; Series</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>Patient Generated Data</strong></td>
<td></td>
</tr>
</tbody>
</table>
“Secondary Use” of Data

• “Reactive” Secondary Use
  – Insurance Claims
    • Single payer
    • Multiple payer
  – Clinical
    • Lab, radiology, pathology, ...
    • EMR, Population Management...

• “Planned” Secondary Use
  – Structured data creation and capture
  – Expanded common data sets outside clinical trials
  – Common intervention protocols

• Patient related and ‘generated’ data
  – Expanded demographics (race, ethnicity, language, SES, ...)
  – Functional status
  – Personal and family preferences and values
  – Genomics, etc...
“Reactive” Secondary Use: Optum Natural History of Disease (NHD) Model

Challenge: Using claims...

– Define a disease protocol
– Find people who match the protocol
– Gather all their medical history
– Find control cohorts for each individual
– Find *their* history
– Compare clinical and financial attributes between the two groups in every conceivable way
– Do this in a minute or two...
“Big Data” and NHD

– Massively parallel database (Netezza)
– Optimized Data Storage structures
– User Interface for mix and match attributes
  • Medical Codes
  • Temporal requirements
  • Member demographics
– SQL based Real Time cohort matching
– Signal Visualizations
Matched sets

Then explore your population financially or clinically. Matched Pairs: 123,021 (246,042 total population)

Target Population

Population with Primary Criteria
1,137,728

Population with Inclusion Criteria
314,996

Population with Exclusion Criteria
N/A

Comparison Population

Matched Pairs: 123,021 (246,042 total population)

292,009
Pre exposure Dx comparison

Confidential property of Optum. Do not distribute or reproduce without express permission from Optum.
“Planned” Secondary Use: Optum Claims Linked Registries (CLRs)

A Registry for Healthcare Operations

– Initial Focus on Breast and Lung Cancer

– The registry is being created for healthcare operations purposes.

– All services provided by Optum in creation of the registry are conducted on behalf of the holder of the data.

– The holder of the data has sole discretion over the design of the registry, use of any PHI and whether to de-identify the data for healthcare operations and/or research purposes.
## Linked Longitudinal Data

### Claims Data

<table>
<thead>
<tr>
<th>Member</th>
<th>Physician &amp; Facility Claims</th>
<th>Other Linkages</th>
<th>Socio-economic Status *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan characteristics</td>
<td>Physician or Facility Identifier</td>
<td>Linked Medical Records</td>
<td>Income</td>
</tr>
<tr>
<td>Age</td>
<td>Procedures (CPT-4, revenue codes, ICD9, J-codes for medications)</td>
<td>• selected relevant clinical data</td>
<td>Net Worth</td>
</tr>
<tr>
<td>Gender</td>
<td>Diagnosis (ICD-9-CM, DRG)</td>
<td>• Fill gap in claims:</td>
<td>Education</td>
</tr>
<tr>
<td>Dates of benefit eligibility</td>
<td>Admission and Discharge Dates</td>
<td>• Identify patients from claims and add clinical data from medical record</td>
<td>Race/ethnicity</td>
</tr>
<tr>
<td></td>
<td>Dates and Places of Service</td>
<td>• Baseline data and annual follow-up</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patient/Plan Paid Amounts</td>
<td>• Periodically add new patients, and follow-up on existing patient</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standardized Costs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Pharmacy Claims

- Prescribing Physician
- Drug Dispensed NDC
- Quantity and Date
- Drug Strength
- Days Supply
- Patient/Plan Paid Amounts
- Standardized Costs

### Other Linkages

- Linked Medical Records
  - selected relevant clinical data
  - Fill gap in claims:
  - Identify patients from claims and add clinical data from medical record
  - Baseline data and annual follow-up
  - Periodically add new patients, and follow-up on existing patient

- Socio-economic Status *
  - Income
  - Net Worth
  - Education
  - Race/ethnicity

* Available for a subset of members.

### Date of Death
A few of the many applicable lines of research...

**In both breast and lung cancer:**
- How many patients have more than one line of therapy? What lines are used? Is there a relationship with survival rates?
- How do survival and costs compare for advanced cancer patients who opt not to have therapy compared to those who are treated?
- Can differences in costs and response rates for various lines of therapy be demonstrated?
- What is the incidence of severe infusion reactions among treated patients? How are they managed and treated? How much do they cost?
- How are chemotherapy regimens altered in response to adverse events (e.g., neutropenia, neuropathy)? Do dose delays or reductions follow?

**Breast cancer:**
- What are the treatment patterns among triple negative (HER2 under expressed, ER/PR negative) breast cancer patients? What are survival rates? Costs? Rates of disease progression?

**Lung cancer:**
- What are the treatment patterns among non-small cell lung cancer (NSCLC) patients? What are survival rates? Costs? Rates of disease progression?
National ‘Secondary Use’ of Data Efforts to Support CER, PCOR and Related Research

• The HMO Research Network
• AHRQ (AcademyHealth): The Electronic Data Methods Forum
• Sentinel (FDA)
• Health Care Cost Institute (HCCI)
• PACES (FDA)
• Multi-Payer Claims Database (ASPE/HHS)
• Registry for Autism Spectrum Disorders (NIMH)
• Others...
Resources for Increasingly Structured Learning
...Examples with a Lewin/Optum Role

Multi-Payer Claims Database (ASPE/HHS)
- Build a database to include multiple data partners (public and private) for CER
  - CMS (MediCare and MediCaid)
  - Commercial Plan
  - Others – possibly state All Payer Databases
- Design to allow both distributed and aggregated analyses
- Beta-test for application to CER
- Develop a strategy for longer term sustainability

Registry of Individuals with Autism Spectrum Disorders (NIMH)
- Develop an advisory group including patients and families
- Build a registry from commercial claims that includes the individuals of interest plus their families
- Validate the registry through chart review
- Test hypotheses about etiology and health services use by patients and families
- Examine options for sustainability
Sustainability...

Data Sources

Data Users

Funding

Governance

Data Sources

Data Users

Funding

Governance
A Framework for Sustainability...
Thanks...

Paul Wallace MD
Director, Center for Comparative Effectiveness Research
The Lewin Group
Paul.Wallace@lewin.com