Using Real-World Data for Outcomes Research and Comparative Effectiveness Studies

Christina Mack, MSPH, PhD
Director, Epidemiology and Health Outcomes, Quintiles
Adjunct Assistant Professor, Epidemiology, UNC-Chapel Hill

Duke/UNC Health Informatics Seminar Series | 11 Nov 2015
Context: Who is Quintiles?

World’s largest Contract Research Organization (CRO)

1,650 Medical Doctors and PhD’s
4,400 Clinical Monitoring Associates
500 Nurse Advisors
5,500 Commercial Sales Representatives
1,500 Project Managers
28,000 Employees Worldwide
Setting Best Practice Standards for Real-World Research

We Help Set the Standards Across Healthcare on how Studies are Performed

- **AHRQ Handbook**
  Best Practice Guidelines for Designing and Implementing Patient Registries

- **AHRQ OCER**
  Standards and Best Practices for Designing Observational Comparative Effectiveness Research

- **GRACE Principles Initiative**
  Good Practice Principles for Observational Comparative Effectiveness Research

- **PCORI**
  Developed Standards in the Conduct of Registry Studies for Patient-Centered Outcomes Research for Inclusion in PCORI Report

We Help Pioneer New Approaches to Data and Studies

- **ROPR** (Registry of Patient Registries)
  A Central Repository for Patient Registries

- **ENCePP Research Center**
  Monitoring of Post-Markedeted Medical Products in Europe

- **EMA PROTECT-EU Partner**
  Innovative Methods in Pharmacoepidemiology & PV

- **ASTER-D**
  A Spontaneous Trigger Approach to Collecting Adverse Event Data through Electronic Data

We Leverage Relationships with Provider Networks, Clinical Foundations, Professional Societies, and Government Agencies to gain exceptional provider insights

*QUINTILES*
Objectives

Pragmatic Approaches to Use of Real-World Data for Research

• Evolving continuum of real world data sources

• Strengths & limitations of different real-world data sources

• Hybrid prospective studies: combining data sources for increased value

• Example applications using real-world data
Types, Strengths, and Limitations of Real-World Data
Background

• Advances have improved and expanded available real-world data sources for outcomes and comparative effectiveness research use
  > Greater leverage of information technology for healthcare data administration
  > Strengthening of observational research methods

• Insights from real-world data have the potential to influence clinical knowledge of treatment effectiveness and safety, value-based purchasing, pricing and access to therapy, and provider and consumer decisions

• Creative approaches to research which consider all available options are key to generating strong results and scientific evidence
Real-World Data Fundamentals

Sources as Defined by ISPOR

Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report

Louis P. Greene Jr, PhD (cochair), Peter J. Neumann, ScD (cochair), Pauline Erickson, PhD,
Deborah Marshall, PhD, C. David Malter, PhD

Sources of RW Data

Real-world data can also be categorized by type of data source. Our Task Force defined six such sources:
1) supplements to traditional registration RCTs; 2) large simple trials (also called practical clinical trials);
3) registries; 4) administrative data; 5) health surveys; and 6) electronic health records (EHRs) and medical chart reviews.

Results: We defined RW data as data used for decision-making that are not collected in conventional randomized controlled trials (RCTs). We considered several characteriz-
Real-World Data Types

*Primary Data*

- **Supplements to traditional RCTs:**
  - Commonly known as trial-based or ‘piggyback’ evaluations
  - Useful for collecting health economic information alongside clinical trials (quality of life, PRO, healthcare utilization)

- **Pragmatic trials:**
  - Commonly known as practical or naturalistic trials; include large simple trials
  - Attempt to measure safety and effectiveness of an intervention in a real-world setting (routine practice)

- **Registries:**
  - Include prospective cohort studies
  - Collect data on group of patients with a given condition or common treatment

All involve a CRF or data collection instrument
Real-World Data Types
Primary Data (Cont.)

• Health surveys:
  › Useful for basic epidemiologic data or macro-level views on utilization
  › Useful for obtaining PROs and patient and physician views

• Medical chart reviews:
  › Abstracting patient demographic and clinical data from patient charts

All involve a CRF or data collection instrument
Real-World Data Types

*Secondary Data*

**Administrative data:**
- Also known as claims data
- Collected for billing purposes
- Organized by bill for service (inpatient, outpatient, physician, Rx)

**Electronic health or medical records:**
- Electronic health records also called electronic medical records (EMRs)
- Aggregated from medical practices, giving point of entry to provider networks
- Collected for patient care purposes
# Classifying Real-World Data Sources

*Two-by-Two Typology Facilitates Critical Review*

<table>
<thead>
<tr>
<th>Primary Data Collection</th>
<th>Retrospective Designs</th>
<th>Prospective Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Medical Chart Review</td>
<td>RCT Piggybacks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pragmatic Trials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Registries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Health Surveys</td>
</tr>
<tr>
<td>Secondary Data Collection</td>
<td>EMR</td>
<td>Automated EMR or claims</td>
</tr>
<tr>
<td></td>
<td>Administrative Claims</td>
<td>data feeds</td>
</tr>
</tbody>
</table>
## Classifying Real-World Data Sources

### Strengths & Limitations of Primary Sources

#### Key Strengths:
- High degree of control over what data are collected—and how
- For stand-alone studies, fine-tuning of sample size is possible
- For stand-alone studies, upfront control of confounding & bias is possible, even for real-world care patterns

#### Key Limitations:
- Prospective data are more costly and take longer to collect than retrospective—sometimes by orders of magnitude
- Sample size may be limited and for RCT piggybacks, no fine-tuning of sample size
- For some diseases, generalizability can be narrowed, as patients enrolled in research may be more engaged in care

<table>
<thead>
<tr>
<th>Primary Data Collection</th>
<th>Retrospective Designs</th>
<th>Prospective Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Medical Chart Review</td>
<td>RCT Piggybacks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pragmatic Trials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Registries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Health Surveys</td>
</tr>
</tbody>
</table>
Classifying Real-World Data Sources

**Strengths & Limitations of Secondary Sources**

**Key Strengths:**

> Data already exist in charts / computer systems ➔ Economy of data collection
> Potential for enormous sample sizes almost instantaneously
> Data reflect real-world patterns of care, not affected by study protocol
> Data mining approaches can uncover key relationships not on clinical radar

**Key Limitations:**

> Data already exist in charts / computer systems ➔ What you see is what you get
> Lack of patients taking new products
> Numerous sources of confounding & bias, not all of which can be controlled
> Claims may not cover prescription; EMR may not reflect continuity of care.
> Coding may vary across physicians and time.

---

<table>
<thead>
<tr>
<th>Secondary Data Collection</th>
<th>Retrospective Designs</th>
<th>Prospective Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMR</td>
<td>Automated EMR or claims data feeds</td>
<td></td>
</tr>
</tbody>
</table>
## RWD Practical Comparison

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Administrative Claims</th>
<th>Electronic Medical Records</th>
<th>Primary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Details</td>
<td>Basic demographics (age, sex) plus enrollment</td>
<td>Demographics plus vital signs, BMI, allergies, smoking status</td>
<td>Flexibility on what is collected (demographic and clinical characteristics and vitals and history)</td>
</tr>
<tr>
<td>Medications</td>
<td>Drug code (name, form, strength), Rx fill date, amt supplied, dose &amp; freq for pharmacy-dispensed drugs; no OTC</td>
<td>Mostly same detail for Rx's written (but no Rx fill date); current meds, including OTC products, available</td>
<td>Detail on medications prescribed including OTC</td>
</tr>
<tr>
<td>Diagnostics</td>
<td>ICD-9 codes</td>
<td>ICD-9 codes, problem lists, severity, symptoms</td>
<td>Full detail from chart with flexibility on collection</td>
</tr>
<tr>
<td>Procedures</td>
<td>CPT® codes</td>
<td>CPT® codes</td>
<td>Full detail from chart with flexibility on collection</td>
</tr>
<tr>
<td>Laboratory</td>
<td>CPT® codes, date; limited availability of lab results</td>
<td>CPT® codes, date, &amp; e-feed of lab results sometimes including pathology &amp; radiology</td>
<td>Detail on labs and pathology collected and results available</td>
</tr>
<tr>
<td>Hospital</td>
<td>Dates of admission &amp; discharge, diagnoses, major procedures; usually nothing on inpatient drugs</td>
<td>Hospital EMR: detail on all aspects of inpatient care, including day:time info; ambulatory EMR: not much</td>
<td>Full detail from chart with flexibility on collection</td>
</tr>
<tr>
<td>Financial</td>
<td>Charges, amounts reimbursed, co-pays</td>
<td>Usually not available</td>
<td>Usually not available</td>
</tr>
</tbody>
</table>
Hybrid Prospective Studies

Combines existing data (e.g., EMR, claims) with primary data (e.g., clinician reported data, PRO, observer reports) to address one or more research questions efficiently and effectively.

Option: For a pragmatic trial, randomize to treatment

- **Primary Data** (Examples)
  - Patient-Reported Data (surveys) / PROs
  - Physician-Reported Data (e.g., direct-to-clinician surveys, COAs)
  - Other Data (e.g., pharmacist-reported)
  - Protocol-defined clinical data

- **Existing Data**
  - Data collected for clinical or administrative purpose (e.g., EMR, claims)

- **Follow-Up Time**
  - Year 1
  - Year 2
  - Year 3
Case Studies
The Research Question Defines Study Design and Data Selection

Five case examples spanning various designs and data sources:

• Retrospective observational study: EMR and claims
• Retrospective observational study: existing registry data linked with insurance claims
• Hybrid prospective study design: EMR and direct-to-patient and provider surveys
• Nested prospective study within an existing registry
• Primary data collection: Prospective patient registry
Comparing Patient-Centered Outcomes after Treatment for Uterine Fibroids

**Sponsor**
Patient-Centered Outcomes Research Institute (PCORI)

**Research challenge**
Guide informed decisions about which uterine fibroid treatment options are most likely to result in outcomes of greatest importance to each patient.

**Objective**
- Compare durability of symptom relief, measured by *subsequent procedures*, after uterus-conserving treatments
- Compare durability of symptom relief, measured by *recurrent symptoms*, after any procedural treatment (including hysterectomy)

Comparing Patient-Centered Outcomes after Treatment for Uterine Fibroids

Key drivers:
- Geographic differences in treatment
- Length of follow-up
- Hysterectomy makes randomization challenging

Durability of Symptom Relief, measured by:
- Subsequent procedures
- Recurrent symptoms
Retrospective Observational Study: EMR and Claims

- **Quintiles EMR and Truven MarketScan claims data**
  > Over 13,000 patients with UF diagnosis
  > Clinical and laboratory information from outpatient EMR records
  > Hospitalizations, emergency dept visits, other medical care and healthcare costs by service type from claims

- **Quintiles COMPASS* Research Network**
  > Over 12,000 patients with UF diagnosis
  > Federated network of integrated healthcare delivery systems (IDNs)
  > Patient information available across the continuum of care; EMR records include hospitals, clinics physician offices

* COMparative effectiveness and PATient Safety and Surveillance
Effectiveness of Chemotherapy in Older Adults

**Sponsor**
Academic

**Research question**
Is FOLFOX (oxaliplatin) more effective than 5-FU in reducing mortality among older stage III colon cancer patients?

**Objectives**
- Understand effectiveness of oxaliplatin vs 5-FU in reducing mortality among older patients with stage III colon cancer, who were underrepresented in clinical trials but are most impacted by this disease
- Develop methods for nonexperimental CER of new therapies including novel propensity score method and investigation of calendar time as instrument for treatment
- Examine impact of calendar time in dynamic settings to understand and account for changes in channeling over time.


Retrospective Observational Study: Registry Data Linked with Insurance Claims

**Key drivers:**
- High number of patients to answer research question
- Required tumor staging and grade, critical variables not found in EMR
- Mortality, treatments and surgeries across continuum of care were critical.

**Data sources:**
- **U.S. cancer registry: Surveillance Epidemiology and End Results (SEER)**
  - Identify variations in cancer incidence, mortality and survival
  - Evaluate impact of advances in prevention, treatment
  - Contains 26% of cancer cases in the U.S.
- **U.S. health insurance claims data: Medicare**
  - Beneficiaries: Disabled, End Stage Renal Disease, Elderly
  - 97% of individuals aged 65+ covered.

Reference: http://healthcaredelivery.cancer.gov/seermedicare/
# Patterns of Diabetes Care

## Sponsor
- Industry

## Research question
How do practice and referral patterns affect outcomes in management of patients with type 2 diabetes?

## Objectives
- Describe usual care practice patterns at the sites providing initial diabetes care overall and by specific site/provider characteristics
- Describe usual care practice patterns and transition of diabetes care at referral sites
- Describe the effect of practice and referral patterns on selected outcomes, including:
  - Glycemic control
  - Titration and Dose
  - Persistence & Adherence
  - Discontinuation & Switching
  - Side effects
  - Complications
Hybrid Prospective Study Design: EMR + Direct-to-Patient and -Provider Surveys

**Key drivers:**

- Reduce burden of data collection on sites
- Important endpoint: reasons for switching and discontinuation

![Diagram showing flow from Quarterly EMR Data to Site Surveys, Provider Surveys, and Patient Surveys (MediGuard), leading to Harmonize data elements within an EDC system, which in turn leads to an Analytic data file.]

[Diagram showing flow from Quarterly EMR Data to Site Surveys, Provider Surveys, and Patient Surveys (MediGuard), leading to Harmonize data elements within an EDC system, which in turn leads to an Analytic data file.]
Treatment patterns and effectiveness for RA

Sponsor

CORRONA Investigators and Genentech

Research challenge

Assess comparative effectiveness of treatment options, safety, and treatment patterns among rheumatoid arthritis patients in the U.S.

Objective

Assess comparative effectiveness and safety of different classes of biologic agents among RA patients initiating either tumor necrosis factor (TNF) antagonists or non-TNF-inhibitor biologic agents

Nested Prospective Study Within an Existing Registry

Use existing registry infrastructure to address a new research question

Key driver:
• Efficient data collection to answer key study goals with high quality data

Study Flow Diagram of CORRONA and CERTAIN

Visit schedule determined by provider, typically every 4 months
Basic CORRONA dataset collected

Patient enrolled in CERTAIN

Patient in CORRONA identified as eligible for CERTAIN; invited to participate; provides informed consent

Mandated visits every 3 months for 1 year
Additional CERTAIN dataset and blood samples collected

Sites and patients return to CORRONA for follow-up to allow studying of long-term comparative safety

Registry in Glaucoma Outcomes Research (RiGOR)

Sponsor
U.S. Agency for Healthcare Research and Quality (AHRQ)

Research challenge
Compare the effectiveness of treatment strategies for primary open-angle glaucoma, in response to the U.S. IOM “Initial National Priorities for CER”

Objective
Compare response to treatment for (1) patients treated with laser surgery and (2) incisional/other surgeries with (3) patients receiving additional medication, at one year post-treatment for the following outcomes:

- $\geq 15\%$ reduction in Intraocular Pressure (IOP) - primary endpoint
- Improvement in Patient-Reported Outcomes and Quality of Life
- Glaucoma severity and visual acuity measures
- Subsequent surgeries, incidence of complications

Primary Data Collection: Prospective Patient Registry

Key drivers:
- Consistency of clinical measures
- Direct to patient measures

RiGOR (NCT01645319/ RoPR ID 3) was able to collect data on a large enough sample to perform subgroup analyses in key populations, and capture endpoints in a real-world setting.
Conclusion

Pragmatic Approaches to Real-World Data Source Selection & Use

• Registries, a rich source of real-world data, are a valuable tool for outcomes research and comparative effectiveness studies

• Other data sources – such as claims data or EHRs – can also be used to address these types of research questions

• Innovative approaches, such as hybrid studies that combine primary with secondary data collection, are an emerging option

• Each type of real-world data has specific strengths and weaknesses

• A systematic approach to selection of data sources is critical
Thank you!

Questions?

Christina DeFilippo Mack, PhD, MSPH
Director, Epidemiology and Health Outcomes
Quintiles Real-World & Late Phase Research
Adjunct Assistant Professor
UNC Department of Epidemiology

Contact: Christina.Mack@quintiles.com