Using Real-World Databases for Evidence Development

Strengths, Weaknesses and Evolving Approaches

Dr Terry Cox
Webinar
14th November 2012
Your Presenter

Terry Cox MD, PhD
Director, Biostatistics at Quintiles Outcome

Terry A. Cox, M.D., Ph.D., is a board-certified ophthalmologist with fellowship training in neuro-ophthalmology and a Ph.D. in biostatistics. Before obtaining his Ph.D. he worked for 18 years as a clinical neuro-ophthalmologist in academic settings, where he had broad clinical and research experience in neurology, ophthalmology, and related disciplines. He obtained a Ph.D. in biostatistics with emphasis in epidemiology at the University of North Carolina at Chapel Hill in 1999 and has worked as a medical statistician at Duke University Medical Center and the National Eye Institute, and has served as statistical consultant for numerous published studies in ophthalmology and neurology. Since joining Quintiles Outcome, he has worked in a wide variety of disease areas, including cardiology, oncology, pediatrics, and diabetes.

Dr. Cox has authored and co-authored numerous research articles in various national and international medical journals. He has received honors and awards in both medicine and statistics, and has been elected to Alpha Omega Alpha (medical honorary society) and Delta Omega (public health honorary society).
Overview

• Global data landscape and emerging data sources
• Different types of real-world data available, including their strengths and weaknesses
• Considerations with using these data for conducting outcomes research
• Evolving complementary research approaches
Today’s Webinar Audience

Academia
Biostatistician
Clinical Operations
Epidemiology
Health Economics/Health Outcomes
Market Access
Medical Affairs
Risk Management
Other
Polling Questions

• A small number of polling questions have been added to today’s webinar to make the session more interactive
Global data landscape and emerging data sources
Polling Questions

• In what regions* have you used real-world databases for research?
  > None
  > USA
  > Canada
  > Europe
  > Asia

*select all that apply
Polling Questions

• Where is your greatest unmet need* for real-world databases for research?
  > USA
  > Canada
  > Europe
  > Asia
  > Latin America

*gap between database demand and availability
Unmet Market Needs

Although real-life data is a growing need, there are a number of issues which confound the collection of data.

Growing Need:

• Payers increasingly asking for more evidence of cost effectiveness that applies to the real world
• Collection of real-life data on the new drug may only be possible after a decision regarding reimbursement
• Modelling accepted by the majority of payers with some providing conditional approval pending the collection of real-world data collected post-launch for validation

Data Collection Challenges:

• A lack of good quality and sufficiently representative databases in many countries:
  – Those that exist often not complete across different health care centers
  – May be focused on GPs or the hospital sector, but rarely cover all the different settings that play a role in medical treatment
  – Often missing data or contain poorly specified information (e.g. on the severity of the condition)
• A further limitation with the description of an event often differing in real-life data compared to a randomized trial
• Particular challenge of prospective data in terms of the manpower effort and budget required to collect it - finding sources that are willing to provide the data can be an issue
The Future is Growth in EMR

- EMR/EHR (electronic medical/health records) is the major segment driving growth in healthcare information technology
- Accenture projects 6.6 to 9.7 percent annual growth in hospital EMR globally
  - North America 9.7 percent
  - Asia Pacific 7.6 percent
  - Europe, Africa, Latin America 6.6 percent
- Global Data projects worldwide EMR growth at 12% annually through 2016
  - Driven by government financial incentives in USA, Australia, China, and Canada
- Recent legislation in USA
  - Affordable Care Act 2010
  - Health Information Technology for Economic and Clinical Health (HITECH) Act 2009
  - Medicare and Medicaid EHR Incentive Programs
EMR Use in Europe

The use of ICT for Health purposes by General Practitioners in Europe varies considerably.

Legend:
- eHealth Frontrunners
- eHealth Average Performers
- eHealth Laggards

Based on scoreboard of electronic storage of patient data, computer use in consultation and electronic transfer of patient data

Source: Benchmarking ICT use among General Practitioners in Europe
Different types of real-world data available, including their strengths and weaknesses
## Post-Approval Study Designs

<table>
<thead>
<tr>
<th>Features</th>
<th>Design</th>
<th>Key Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT (explanatory)</td>
<td>Prospective Intervenional</td>
<td>‘Gold standard’, randomized, blinded, placebo controlled to measure efficacy in highly selected patients with standardized and intense follow up</td>
</tr>
<tr>
<td>RCT (pragmatic)</td>
<td>Prospective Intervenional</td>
<td>Randomized, open-label, with ‘usual care’ control group to measure effectiveness in broad patient population with follow up</td>
</tr>
<tr>
<td>Observational</td>
<td>Prospective Observational</td>
<td>Non-randomized non-interventional study that monitors a cohort(s) over time</td>
</tr>
<tr>
<td>Registry</td>
<td>Prospective Observational</td>
<td>Observational study of patients with a specific disease or receiving a specific treatment</td>
</tr>
<tr>
<td>Database Analysis</td>
<td>Retrospective</td>
<td>Capture patient data from existing database. *Increasingly used for prospective research</td>
</tr>
<tr>
<td>Chart Review</td>
<td>Retrospective</td>
<td>Capture patient data from patient record</td>
</tr>
<tr>
<td>Model</td>
<td>Retrospective</td>
<td>Decision analytic framework that simulates the impact of intervention using multiple data sources</td>
</tr>
</tbody>
</table>
Design Strengths and Weaknesses

Strength of Evidence
- Stronger
- Weaker

Type of Study
- RCT (explanatory)
- RCT (pragmatic)
- Observational
- Registry
- Database Analysis
- Chart Review
- Decision Model

Purpose
- Efficacy
  - Can it work?

Validity
- Internal
- External

Effectiveness
- Does it work in the ‘real world’?
Why not use RCTs for all research needs?

- RCTs are not well suited to answer all research questions.
  > Atypical behavior & setting
    - Protocol-driven behavior in narrowly defined study population
    - Not usual physician or usual practice
  > May be analyzed by intent-to-treat, not how drugs are actually used
  > Do not give insights into why clinicians may use drugs off-label or in risky situations
- Generally small subject numbers
- Short duration of treatment and/or follow-up
- May use surrogate/intermediate endpoints
- Comparator may be placebo, specific comparison with a single alternative, or “standard of care”
- Poses special challenges for rapidly changing products like medical devices
- Masking/blinding is almost impossible in many device studies
Advantages of Studies using Databases

- Less expensive and time-consuming
- Limited inclusion/exclusion criteria mean patients are more representative of usual practice
- **Observe practice not dictated by protocol**
- **Comparative information from actual practice**
- Estimates of treatment impact are more realistic
- Off-label use
- Can be conducted in situations where RCT are impractical or not feasible
  - Compliance and adherence
  - Follow-up for long-term benefits or delayed complications under conditions of real-world use
  - Heterogeneous patient populations, including special patient subgroups
  - Events that occur with low frequency (condition or outcome)
  - Safety studies
Uses of Real-World Databases

Natural history of disease and current practice patterns

- **Can be used to design better clinical trials**
- Describe and quantify population that may benefit from a new drug in development
  - Number of potential patients?
  - Co-existing illnesses (co-morbidities)
  - Impact of the disease?
    - E.g., Mortality, functional impairment, independent living, absenteeism, presenteeism, etc.
- Effectiveness of currently available treatments?
- Obtain population-based rates of adverse events with previous, similar drugs
- Factors determining treatment choice
Real-World Databases in the Product Continuum

Provide real-world data on:

- natural history of disease
- burden of illness
- treatment patterns
- competitor products and
disease management

to inform development, launch strategy, and market access (product not included).

Provide real-world data on:

- brand usage (on and off-label)
  - safety
  - effectiveness
  - compliance, adherence, persistence
  - treatment satisfaction
- competitor brands
  - comparative effectiveness
- disease management
Data Generalizability

Heterogeneous General Population

- Older patients with multiple comorbid diseases and meds
- Younger patients with multiple comorbid diseases
- Older, otherwise healthy
- Younger, otherwise healthy

Database Sampling
(few or no exclusion criteria leading to a proportional representative study population)

RCT Sampling
(multiple exclusion criteria leading to a homogenous study population)
Database Analysis is best when:

- Study results need to be generalizable
- Need answers/insights on a large number of patients quickly
- Need estimates of drug utilization and health care resources in routine clinical practice
- Need to identify trends over time
- The outcomes or disease of interest are rare events
- Randomization is an issue because of ethical considerations
- It is challenging to recruit the target patient population, e.g., patients with depressive illness or patients with rare characteristics
- The research budget is limited and it is not feasible to conduct a resource-intensive study (e.g., site management costs)
Considerations with using these data for conducting outcomes research
Polling Questions

• What is your database research experience?
  > None
  > EMR
  > Claims
  > EMR & Claims
EMR vs. Claims Data

- **EMR**
  - Hospital
  - Ambulatory
  - Diagnosis: Problem list, medications, billing codes, labs

- **Claims**
  - Hospital
  - Ambulatory
  - Pharmacy
  - Insurance
  - Diagnosis: Billing codes
EMR vs. Claims Data

**EMR Database**
- Updated frequently
- Lab data available for majority
- Less population churn
- Longer active history
- Codification: concept standard
- Narrow but deep
  - Extensive unstructured clinically rich data

**CLAIMS Data**
- Time lag of 3-6+ months
- No linking of Rx to Dx
- Insurer specific: turnover
- Variable enrollment period
- Standardized cleaning
- Broad but less detail
  - Reimbursement data but not clinical richness
Polling Questions

• What are your greatest challenges in using real-world databases for research?
  > No database for research question
  > Data access
  > Data quality
  > Data completeness
  > Other
General Challenges for Database Use

- Database may not be available for the research questions/objectives
- Incomplete data: consider preliminary study of data availability
- Free-form text entry
- Medications prescribed, not prescriptions filled
- External test results may not be available
- Data from referred patients may be limited
- Data not available from both hospital and ambulatory settings
- Problem lists may not be comprehensive: chronic diseases may not be listed at every visit
  - Patients with new diagnosis in EMR may have long-standing disease
  - Disease duration may be difficult to assess
- Limited source data verification
Specific Challenges for Database Use

• Assignment of treatments is shaped by clinical judgment and patient preference
  > Nothing random about it
  > Leads to systematic differences between patients who receive different treatments

• Key challenge for analysis and interpretation is to “control” for known & unknown differences between groups
  > Statistical methods: propensity score matching, instrumental variable analysis

• Other sources of bias:
  > Non-random sample
  > Misclassification, including incorrect diagnoses
  > Missing data, including loss to follow-up
Design of Database Studies

- Internal comparison group
  - Cohort
    - Prospective
    - Retrospective
  - Case-control
    - Prospective
    - Retrospective

- External comparison group
  - Historical
  - Other databases
  - RCTs
  - Observational studies

- No comparison group
  - Some safety studies
  - Descriptive analyses (treatment patterns, resource utilization, etc.)
Sample Size Considerations

- Small samples generally not a concern
- Everything is significant in large datasets
  > Clinical vs. statistical significance
- More patients can be more expensive, but less so than in RCTs
- Estimates based on precision (external/no comparison group)
  > Example: What proportion of patients show a treatment response?
  > Estimate will have confidence interval, and width of confidence interval determined by sample size.
- Estimates based on internal comparison group
  > Similar to sample size estimation in RCTs
  > Adjustments for propensity score matching, missing data, etc.
Evolving complementary research approaches
Polling Questions

• In comparison to others, do you feel that your company is innovative in the use of real-world databases?
  > Less innovative than other companies
  > About the same as other companies
  > More innovative than other companies
A Novel Hybrid Study Design

- Longitudinal, observational study in physician office settings with EMR to assess practice patterns, patient experiences, and outcomes
- Multiple EMR networks
- Automated/passive data collection through EMR + active data collection through surveys
- Retrospective + prospective

1 year

3 years
A Novel Hybrid Study Design: Implementation

- Sites identified by EMR vendor
- Patients identified in EMR database, then invited to participate
- Online registration and consent
- Baseline patient, provider, and site surveys
- Periodic prospective online patient surveys
- Concomitant information from prospective health care provider surveys
- Continual assessment of EMR for new qualifying patients
A Novel Hybrid Study Design: Advantages

• Reduces research burden at the site
• Cross-validation of PRO, physician, and EMR data
  > Duration of disease
  > Medication history
• One-year retrospective look-back through EMR appended to three-year prospective follow-up simulates a four-year study
Hybrid Observational Research Models

- 80% of adults are seeking health information on-line

- This rapid adoption of the internet coupled with technology advances has created research opportunities
  > Rapid access to large numbers of patients seeking health information
  > Cost-effective, bi-directional communication via email and SMS

- Hybrid observational research models can include PRO and EMR
  > Some studies recruit via physicians, while others approach patients directly (direct-to-patient studies)

Wales Cholesterol PRO+EMR

**Background**

- **Objective**: Build a UK data-rich environment to measure outcomes
- **Approach**: Conducted study to demonstrate PRO+EMR link with SAIL data warehouse in Wales, UK
  - In 6 weeks, recruited 240 cholesterol patients who completed PRO assessments and consented to share identifiers
  - Identifiers provided to NHS Wales to create pseudo-identifier bridge into SAIL data warehouse
    - PRO data combined with electronic data stored in SAIL at Swansea Univ. for all Wales

**Findings**

- **% Match from PRO to EMR Wales, UK**
  - Patients (224 of 240): 93%
  - Diagnosis (89 of 91): 98%

**Conclusions**

- PRO+EMR process proven feasible in the UK
- Process found to be transferable to Scotland

US Case Study Example: PRO+EMR+Lab

Background

• **Objective:** Identify genetic factors associated with biologic response

• **Approach:**
  - Enroll 1,000 patients who self-report biologic exposure
  - Capture consent, deliver questionnaire, and obtain electronic and paper signature for medical record release
  - Subjects to be mailed DNA sample collection kit and addressed, postage-paid envelope for sample return
  - All data entered into dataset based on Study ID (de-identified)

Process

PRO
• Informed consent
• Medical History
• RADAI, WPAI, MARS-5

Chart
• Electronic record release
• Paper record release

Lab
• DNA sample collection kit

Results

• FPI – July 2012
• 500 subjects recruited in first 6 weeks
Conclusions

• Real-world databases will be increasingly available globally in the next decade.
• Database analysis supplements RCT data to provide a more comprehensive picture of efficacy and effectiveness.
• Real-world databases can accommodate a number of study designs.
• Real-world databases present unique challenges as well as opportunities for research.
Upcoming Events

Post-Approval Summit

• May 7-8, 2013
• Conference Center at Harvard Medical School, Boston, MA
• Key Topics:
  > Comprehensive Approaches to Evidence Development for Safety and Effectiveness
  > Evolving Roles of the RCT and Observational Research
  > Big Data: Leveraging EHR and Health System Data for Safety and Effectiveness
  > Updates on Changing Safety and Risk Management Requirements
  > Comparative Effectiveness, Market Access and HTA
  > Approaches and Models for Addressing Multi-Stakeholder Demands