Design and Analytical Considerations in Registry-Based Studies: Registries 201

Aaron Mendelsohn, PhD, MPH
GRP Webinar | 14 February 2012
Outcome, Cambridge, MA USA & St-Prex, Switzerland
Agenda

• Introduction/Registries defined
• Design options in patient registries
  – Selection of controls
  – Inception cohorts
• Types of data and data collection
• Analytic considerations with registry data
• Wrap-up, Q&A
Introduction / Registries Defined
What is a Patient Registry?

A patient registry...

- Is an organized system that uses observational study methods to collect uniform data (clinical and other)
- Evaluates specified outcomes for a population defined by a particular disease, condition, or exposure
- Serves a predetermined scientific, clinical, or policy purpose

Typical Goals of Patient Registries

**Effectiveness**
- Evaluate clinical or comparative effectiveness

**Safety**
- Measure or monitor safety and tolerability, including comparative risk-benefit

**Quality**
- Measure and/or improve quality of care

**Natural history**
- Incidence and prevalence
- Trends
- Identification of high risk groups
- Resource utilization
- Paths to diagnosis and treatment, etc.

*Goals are not mutually exclusive*
Types of Registries

• **Product**
  - Pharmaceutical / biotech product registries
  - Device registries
  - Pregnancy registries (exposed population = fetus)

• **Disease or Event**
  - Acute disease or event
  - Chronic disease
  - Rare disease

• **Health care service, procedure or clinical encounter**
  - Procedure or hospitalization registries
  - Clinical service and quality measurement registries
# Registries are Real-World Studies

Compared to Trials, Real-World Studies Offer...

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reality</strong></td>
<td>Real-world practice and outcomes</td>
</tr>
<tr>
<td><strong>Applicability</strong></td>
<td>Physician practice and resulting outcomes of that behavior</td>
</tr>
<tr>
<td><strong>Generalizability (external validity)</strong></td>
<td>Broad inclusion/limited exclusion criteria resulting in diverse study populations, often including many subgroups not traditionally studied in RCT</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td>Limited number of RCTs relative to the number of decisions that need to be made</td>
</tr>
</tbody>
</table>
Design and Use of Patient Registries
What Drives Registry Quality and Success?

• Thoughtful planning
  – Engaging stakeholders
  – Choosing optimal study design

• Efficient and seamless operations

• Sound scientific analyses
  – Addressing and minimizing bias
  – Appropriate statistical methodologies
  – Recognizing limitations of data

• Strategic communication of results
Registry Design: Key Considerations

- Formulate a research question(s)
- Identify available resources (sites, clinicians, patients)
- Translate questions of clinical interest into measurable exposures and outcomes
- Choose appropriate study design
- Determine sources of data
- Select patients for study
  - Comparison group? Sample size?
- Determine duration of follow-up
- Assess threats to internal and external validity
Comparators in Registries

• Comparators are most useful in registries where it is important to:
  – Distinguish between alternatives
  – Assess differences, or magnitude of differences

• Selecting a comparator is much more challenging in registry-based studies than RCTs
  – Registries must use study design options and analytical strategies to ensure comparability between groups
Comparator choice directly drives confounding

 Comparator choice directly impacts results

- Low dose comparator → Study drug more effective
- High dose comparator → Study drug safer

Courtesy of Tobias Gerhard, Rutgers U
## Types of Comparators in Registries: Internal Comparators

<table>
<thead>
<tr>
<th>Description</th>
<th>Strengths</th>
<th>Weaknesses, Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Same data source as patients of interest</td>
<td>• Controls are more likely to be like “cases”</td>
<td>• Not always possible to have an internal comparator (e.g., not feasible to have patients not exposed to product)</td>
</tr>
<tr>
<td>• Do not have condition or exposure under investigation</td>
<td>• Consistency in measurement of data</td>
<td>• Expense in following additional patients</td>
</tr>
<tr>
<td></td>
<td>• Ideal for treatment patterns that change over time</td>
<td></td>
</tr>
</tbody>
</table>

e.g., arthritis patients using acetaminophen vs. those taking another product for pain management
## Types of Comparators in Registries: External Comparators

<table>
<thead>
<tr>
<th>Description</th>
<th>Strengths</th>
<th>Weaknesses, Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Different data source as patients of interest</td>
<td>- Efficiency in using existing information</td>
<td>- Similarities in data elements of interest</td>
</tr>
<tr>
<td>- Do not have condition or exposure under investigation</td>
<td>- Not susceptible to changes in treatment practices over time</td>
<td>- Logistics in merging data, privacy issues</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Differences in underlying populations</td>
</tr>
</tbody>
</table>

*Surveillance Epidemiology and End Results*

- SEER data for registries with cancer outcomes, provides benchmarks
### Types of Comparators in Registries: Historical Comparators

<table>
<thead>
<tr>
<th>Description</th>
<th>Strengths</th>
<th>Weaknesses, Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Do not have condition or exposure under investigation</td>
<td>• Efficiency in using existing information</td>
<td>• Susceptible to bias by changes over time (e.g., practice pattern changes)</td>
</tr>
<tr>
<td>• Patients for whom data were collected in the past</td>
<td>• Concurrent control group (internal or external) may not be possible</td>
<td>• Similarities in data elements of interest</td>
</tr>
<tr>
<td></td>
<td>• Appropriate when older treatments are not, or are rarely, used</td>
<td>• Logistics in merging data, privacy issues</td>
</tr>
<tr>
<td></td>
<td>• Hawthorne effect</td>
<td>• Differences in underlying populations</td>
</tr>
</tbody>
</table>

Historical comparators are often used for pregnancy-based studies given the large body of population-based surveillance data, e.g., Metropolitan Atlanta Congenital Defects Program.
When are Exposed Only Populations Used?

- Goal of registry is primarily descriptive or hypothesis generating
  - e.g., describe long-term safety
- More clearly elucidate safety profile
  - Better characterize observed adverse events
- Diseases/conditions for which there is not a viable tx alternative, such as vaccines
- Regulatory purposes
  - Assess the effectiveness of risk mitigation strategies
  - Examine the characteristics of patients prescribed the product of interest and off-label use
# Exposed Only Population Examples

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Registry Name, Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive study</td>
<td><strong>Norditropin Registry of Growth Hormone Patients</strong> to better characterize safety, esp. long-term risk, in patients receiving product</td>
</tr>
</tbody>
</table>
| Further elucidate safety profile, specific safety signals | 1. **Ellidel® (pimecrolimus) Registry in Children** with primary goal of examining rate of lymphoma in children aged 2-17 years exposed to product. Secondary outcomes include examining rate of thyroid cancer and to compare the number of systemic malignancies to SEER data.  
2. **Forteo® (teriparatide) Patient Registry** to examine long-term safety associated with the drug (anabolic tx for osteoporosis), esp. regarding the risk of osteosarcoma |
| No alternative therapy exists | **Palivizumab Outcomes Registry**, operating from 2000 – 2004 and involving 19,000+ infants to describe use of product and hospitalizations RSV-related). Study found that product usage was consistent with AAP guidelines. |
## Exposed Only Population Examples (2)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Registry Name, Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Management Activities</td>
<td><strong>1. Teva Clozapine Patient Registry</strong>, part of a REMS, to prevent clozapine rechallenge in patients at risk of clozapine-induced agranulocytosis. Laboratory data, including WBC and ANC, must be provided and above minimal levels for product to be dispenses. Registry currently includes data on 49,000+ pts.</td>
</tr>
<tr>
<td></td>
<td><strong>2. Isotretinoin Registry</strong>, part of a REMS (iPledge program), to ensure that all pts prescribed product meet eligibility criteria and monthly program requirements to address teratogenic effects of isotretinoin.</td>
</tr>
<tr>
<td></td>
<td><strong>3. Tysabri® Registry</strong>, part of overall REMS, to determine incidence and risk factors in exposed patients for progressive multifocal leukoencephalopathy (PML).</td>
</tr>
</tbody>
</table>

---

Proprietary and Confidential

OUTCOME®  
A Quintiles Company
Many scenarios for which an exposed only population may be used in a registry-based study

While there is utility in exposed only population studies, the major challenge involves interpretation of adverse events and whether observed effects are due to the intervention or an underlying characteristic of the study population

– e.g., cardiovascular events among men in a registry-based product registry for alopecia
New User Designs/Inception Cohorts

- Identifies all patients in a defined population who initiate treatment
  - Usually requires a washout period to identify initial exposure

- Population is followed prospectively from baseline (date of drug initiation)
  - Ideal for comparative studies, approach is similar to RCTs

- Advantages of new user design:
  - Population reflects all users of product
  - Longitudinal experience of users is captured and ascertainment of their experience will be comparable
Biases with Existing Product Users

Under-ascertainment of early events in prevalent users

- Increased early risk with some treatments and interventions (surgeries):
  - Venous thromboembolism risk in women taking third generation OC drugs relative to earlier products
  - Falls after initiating benzodiazepines
  - NSAIDs and peptic ulcers
  - Intussusception with rotavirus vaccine

From: Ray, AJE, 2003
Biases with Existing Product Users (2)

- **Adherence bias:** adherence may be a marker for several unmeasured factors associated with better prognosis

- **Inability to control for disease risk factors that may be altered by the study drugs** (i.e., on causal pathway)
  - e.g., psychotropic meds and falls
    - Falls in prevalent users may be due to psychomotor effects of the drugs (arguing against control);
    - Alternatively, such users may have somatic impairment or major depression that increase risk of falls (arguing for control)

- **“Healthy” cohort effects:**
  - Those with optimal experiences continue to use given drugs and be available for continued study
    - Potential for survivorship bias, lost-to-follow up bias
Weaknesses of New-User Designs

• **Sample size and patient population issues**
  – Loss of prevalent cases, decrease in statistical power
  – Does not permit examination of long-term risk without existing cases
  – New users will disproportionately include more poor compliers and those who do not respond well to the given product

• **Logistical considerations**
  – Identifying new users (washout period)
  – Collecting information on potential confounders at baseline and properly controlling for confounding
What is Channeling Bias?

- Channeling bias, selective prescribing, or confounding by indication/confounding by severity is a form of selection bias where drugs with similar therapeutic indications are prescribed to groups of patients with prognostic differences.
  - e.g. sicker patients being more likely to receive a new drug
Examples of Channeling Bias: Asthma

Confounding by Indication and Channeling over Time: The Risks of $\beta_2$-Agonists

Lucie Blais,1,2 Pierre Ernst,1,3 and Samy Suissa1,2

A previously published nested case-control study, the Saskatchewan Asthma Study spanning 1980–1987, investigated the risk of fatal or near-fatal asthma from various $\beta_2$-agonists, fenoterol and salbutamol. The authors observed confounding by indication because of channeling of inhalers in these studies. Their study used matched subcohorts selected from a cohort of Saskatchewan residents. Using three computerized databases of Saskatchewan Health and following up on hospital admissions for asthma, they measured whether greater asthma severity was associated with preferential prescribing of a first prescription of fenoterol over salbutamol, and whether they were associated with the choice of fenoterol over salbutamol in sequential exposure to inhaled bronchodilator agents. They found an increased risk of hospital admission for asthma in patients prescribed fenoterol only, and a decreased risk in patients prescribed salbutamol only. The reason for this was, however, minimally related to asthma severity. They conclude that the use of fenoterol and salbutamol in the SAEP may have been biased by preferential prescribing. They advocate that long-term information on medication use is essential to ensure that control studies are not biased by indication. Am J Epidemiol 1996;144:1161–9.

Severe asthmatics were most likely to be switched to inhaled fenoterol from salbutamol, thus creating the appearance of a worse safety profile for fenoterol.
Examples of Channeling Bias: COPD

Inhaled Corticosteroids and the Risk of Mortality and Readmission In Elderly Patients with Chronic Obstructive Pulmonary Disease

DON D. SIN and JACK V. TU

The Institute for Clinical Evaluative Sciences (ICES) and The Department of Medicine, Section of Epidemiology, University of Toronto, Toronto, Ontario; and Department of Medicine, University of British Columbia and Vancouver Coastal Health, Vancouver, BC.

Use of oral corticosteroids was associated with an increase in mortality and rehospitalization...patients with increased COPD severity receive these medications...

<table>
<thead>
<tr>
<th>Medication</th>
<th>Mortality (95% CI)</th>
<th>Readmission (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral corticosteroids</td>
<td>1.37 (1.25 to 1.50)</td>
<td>2.09 (1.97 to 2.20)</td>
</tr>
<tr>
<td>Anti-infectives</td>
<td>1.08 (0.99 to 1.17)</td>
<td>1.17 (1.10 to 1.23)</td>
</tr>
<tr>
<td>Antihistamines</td>
<td>1.01 (0.90 to 1.12)</td>
<td>1.20 (1.2 to 1.27)</td>
</tr>
<tr>
<td>Statins</td>
<td>1.00 (0.90 to 1.11)</td>
<td>1.02 (0.96 to 1.10)</td>
</tr>
<tr>
<td>ACE inhibitors</td>
<td>1.00 (0.90 to 1.12)</td>
<td>1.00 (0.97 to 1.01)</td>
</tr>
<tr>
<td>Beta blockers</td>
<td>0.71 (0.65 to 0.78)</td>
<td>0.76 (0.71 to 0.80)</td>
</tr>
</tbody>
</table>

OUTCOME

A Quintiles Company
Gold salts, indicated to delay RA development... preferentially prescribed to patients with advanced disease and worse prognoses
Examples of Channeling Bias: MS

- New oral disease modifying treatment (DMT) for relapsing-remitting multiple sclerosis

DMT competitors are injectable interferon beta, administered IM weekly or SC 3x/week

Patients switching to oral products may be sicker and/or less adherent
Understanding Types of Data and Analyses
Data Collection

• **Primary**: Data collected directly from physicians, patients
  – Flexibility in data available for analysis (e.g., PROs)

• **Secondary**: Data obtained via existing resources
  – Secondary data collection is efficient, less costly
  – Requires identification of appropriate data source

• Study design and objectives will influence the method of data collection

• Data may come from multiple sources, including both primary and secondary sources
Sources of Secondary Data Collection

- **Chart Review/Health records/EMR**: data collected through routine care
  - Includes sociodemographics, clinical and laboratory data, medications/tx, (possibly) behavioral risk factors

- **Administrative/claims data/pharmacy (PBM) data**:
  - Based upon reimbursement for care
  - Data dependent upon quality and nuance of coding
Selecting Specific Data Elements

- Each data element should support registry and address a specific scientific need or question
- Distinguish between critical and “nice-to-know” data
- Engage stakeholders and relevant experts
- Use clinical data standards whenever possible
  - e.g., ACC for acute coronary syndromes, NCI for common cancer data elements
- Determine when data elements will be collected (longitudinal analyses)
  - Group related variables (e.g., dietary information and a fasting blood sample for lipids)
- Pilot test whenever possible
# Provider- and Patient-Reported Data

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Strengths &amp; Uses</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider-reported</td>
<td>• More specific and consistent information than available through coded data or medical record</td>
<td>• Clinicians are highly sensitive to burden</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consistency in capture of certain variables, esp. patient symptoms and use of non-prescribed therapy</td>
</tr>
</tbody>
</table>
| Patient-reported   | • Obtain information on treatments not necessarily prescribed by clinicians (e.g., OTCs, herbal meds)  
• Obtain compliance information  
• Useful when timing of follow up is not concordant with timing of clinical encounter | • Literacy, language barriers  
• Need for validated instruments  
• Loss to follow up, non-participation  
• Ability to report clinical and healthcare utilization information |
Patient-Reported Outcomes (PROs)

- Increased attention to PROs and patient-centered health care system
  - FDA, IOM, PCORI
- Typically for quantifying health status
  - Symptoms
  - Extent to which disease limits patients physically, emotionally, and socially
  - Quality of life
- PROs may be more sensitive to clinical changes than objective measures
- Instruments may involve generic health measures, disease-specific, health utilities
### Attributes of Health Instruments

#### Table 7: Key Attributes of a Health Status Instrument

<table>
<thead>
<tr>
<th>Measurement Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity</td>
<td>The measure quantifies what it is intended to</td>
</tr>
<tr>
<td>Reliability</td>
<td>Reproducible results are obtained when repeatedly given to stable patients</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>The measure is sensitive to clinical change</td>
</tr>
<tr>
<td>Interpretability</td>
<td>A clinical framework is available to interpret cross-sectional data and changes in scores</td>
</tr>
<tr>
<td>Translations exist</td>
<td>Linguistically and culturally appropriate translations are available</td>
</tr>
</tbody>
</table>
Data Analysis

• Specific methodology will depend upon several factors
  – Cross-sectional vs. longitudinal
  – Potential biases and confounding
    • What biases may be present
    • Magnitude of biases
  – How exposure, outcome are defined

• Develop a Statistical Analysis Plan (SAP)
• Conduct sensitivity analyses
Interpretation of Registry Data

Are the results generalizable, valid, and reliable?

1. Representativeness of patients and clinicians?
   - Study inclusion/exclusion criteria
   - Clinical settings used

2. Validity of conclusions?
   - Bias
     - Confounding
     - Selection bias
     - Information bias

3. Reliable results?
**Definition:** Distortion of the exposure-disease association due to extraneous factor(s)

**Necessary conditions for confounding:**

1. Must be a risk factor for the disease
2. Must be associated with the exposure
3. Must **not** be a consequence of exposure/intermediate variable
Confounding example

Consumption of coffee associated with risk of pancreatic cancer
Selection Bias

• Systematic error from differences in characteristics between those studied and not studied

• Relates to generalizability of the study population
  – Sampling frame should be representative of the target population

• Participation rate needs to be evaluated, along with inclusion/exclusion criteria
  – Differences between participants and non-participants
  – How broad are inclusion/exclusion criteria
Information Bias

• Systematic error related to **measurement or classification** of study participants according to one or more variables (risk factors/treatments or outcomes/disease states)

• **Considerations when classifying exposure/treatment**
  – **Use**: ever/never, past/current
  – **Dose/duration**: cumulative duration vs. average dose
  – **Pattern of use**: continuous vs. episodic
    • Time-dependent exposures

• **Consideration when classifying outcome**
  – False positives and negatives
    • Especially problematic in electronic DB studies
    • Should validate disease/non-disease status
Addressing Confounding and Bias in Registries

• Design Phase
  – Restriction/narrowing of inclusion, exclusion criteria
  – Variations in study design, data collection

• Analysis Phase
  – Stratification
  – Multivariable modeling
  – Time-dependent covariates/methodologies
  – Propensity scores
Wrap-up and General Advice in Registry Design and Analysis
Validity Assessment

• **Internal**
  - Data checks using range and logical consistency checks
  - Relevant information for key exposures, risk factors, confounders
  - Compare sample with another source for outcome and exposure
  - Loss-to-follow-up considered, compare completers to drop-outs
  - Duration of follow-up sufficient to capture main outcomes

• **External**
  - Broad inclusion and limited exclusion criteria
  - Assess selection bias, compare analytical to target population
General Advice on the Design and Analysis of Registry-Based Studies

- Engage stakeholders and develop operational and strategic plans
- Build upon existing resources and infrastructure whenever possible
- Avoid a laundry list of objectives/endpoints
  - Prioritize
  - Clear link from objectives to data collection to analysis
  - Make endpoints specific, relevant, and measureable
- Avoid “Death by Data”
  - Balance maximizing data collection with minimizing site and patient burden
  - Use PROs appropriately
  - Focus on “must have” data
- Appropriate study design and methodology will ensure completion of study’s objectives and minimize bias
Closing Thought: Successful Registries Generate and Deliver Value to All Partners

- Industry/Sponsors
  - meet commitments
  - demonstrate safety
  - prove value
  - secure reimbursement
  - generate publications
  - create relationships

- Physicians
  - obtain evidence
  - advance science
  - improve care
  - ensure reimbursement

- Regulators
  - detect safety signals
  - ensure long-term effectiveness and outcomes

- Payers
  - determine value and coverage
  - monitor usage within criteria

- Patients
  - my own health - what choices do I have?
  - what are the risks/benefits?
Contact Information

Aaron Mendelsohn, Director of Epidemiology
+1 (301) 310-1281  amendelsohn@outcome.com

Outcome, a Quintiles Company
201 Broadway, Cambridge, MA 02138 USA
Ch. du Glapin 6, 1162 St-Prex, Switzerland
Upcoming Presentations

  - March 21, 2012
  - [https://www1.gotomeeting.com/register/145722929](https://www1.gotomeeting.com/register/145722929)

- **Webinar**: Experimental vs Non-Experimental - The Strengths and Limitations of Different Methods for Conducting Comparative Effectiveness Research
  - April 2012
  - Registration details coming soon

- **Summit**: Post-Approval Summit at Harvard Medical in Boston, MA
  - May 1-2, 2012
  - [www.postapproval.org](http://www.postapproval.org)

- **Summit**: European Post-Approval Summit in Zurich, Switzerland
  - September 11-12, 2012
  - [www.europe.postapproval.org](http://www.europe.postapproval.org)