Module 7a: Planning Analyses of Observational Data

Category 3: Data Integrity and Rigorous Analyses

Prepared by Jodi Segal, MD, MPH
Scott Zeger, PhD

Presented by Jodi Segal, MD, MPH
Key Considerations

- What is the study population?
- What are the exposures of interest?
- What are the outcomes of interest?

Certainly you will need to define key confounders, but you must start with items above
Common Study Designs Often Used With Secondary Data

- Retrospective study (cohort)
- Case-control study
- Case-crossover design
- Pre-post design
- Interrupted time series
Retrospective Cohort

Target population

Study population

Exposed

Unexposed

observer

losses
outcome

time
Retrospective Cohort

Target population

Study population

Exposed

Differently exposed

time

losses

outcome

observer
Who’s in and who’s out

Probably some demographic characteristics

Possibly a condition of interest

Certainly *exposures of interest* in a cohort study
Masica, et al. (2013): Defining the Study Population (Including the Exposures)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis status</td>
<td>New diagnosis of diabetes captured in the electronic health record (EHR) between 1 January 1998, and 31 March 2009</td>
</tr>
<tr>
<td>Timing of diagnosis</td>
<td>Entry of diabetes diagnosis in the EHR problem list occurring ≥30 days after the first recorded EHR office visit</td>
</tr>
<tr>
<td>Validation as a new diabetes case</td>
<td>Cases identified as new in the EHR underwent a second classification step using external billing data.</td>
</tr>
<tr>
<td></td>
<td>Patients with a diabetes-related bill with a date of service ≥ 90 days before EHR problem onset date were designated as pre-existing and excluded from the cohort</td>
</tr>
<tr>
<td>Age</td>
<td>≥ 21 years at the time of diabetes diagnosis</td>
</tr>
<tr>
<td>Exposure to diabetes medicines</td>
<td>Minimum of 90 days of exposure to metformin, sulfonylureas, or thiazolidinediones (or any combination of those three agents) over the study period, and</td>
</tr>
<tr>
<td></td>
<td>Absence of any documented history of insulin use prior to OAD</td>
</tr>
</tbody>
</table>

“The initial OAD [oral antidiabetic drug] regimen was designated to be the first drug(s) in which a patient had ≥90 days of exposure. ... Exposure times to OAD regimens were cumulative but stopped with a medication prescription gap of ≥90 days or if the medication was switched to another class. In cases where multiple OADs were used concurrently, exposure time was counted independently for each OAD rather than grouped as a combination. For individual patients receiving different OAD monotherapy over time (e.g., sulfonylureas only for 1 year followed by 2 years of metformin only), those exposures were also counted independently in the analytic models.”
“For the proteinuria endpoint (defined as microalbuminuria or worse), we required that patients have a negative urine protein test at first urine measurement following diabetes diagnosis that subsequently became positive during follow-up. Proteinuria tests were interpreted as positive or negative based on the scheme in the Appendix.

The renal function decline endpoint was reached if a patient with an eGFR ≥60 ml/min/1.73 m² at first measurement had an ensuing eGFR that was <60 ml/min/1.73 m² during follow-up. Laboratory results required to evaluate outcomes were obtained directly from the EHR…”

Case-Control Design

No outcome

Outcome

Target population

Study population

exposure

time

observer
Defining (and Describing) the Study Population

- Who’s in and who’s out
- Probably some demographic characteristics
- Possibly a condition of interest
- Certainly *outcomes of interest* in a case-control study
“Acute Pancreatitis Case Patients—We identified patients with acute pancreatitis from the National Patient Registry, which contains data on hospital inpatient discharges from all nonpsychiatric hospitals in Denmark since 1977 and on emergency department and hospital outpatient clinic visits since 1995. Each hospital contact is associated with one primary diagnosis (the one listed first) and up to 20 secondary diagnoses, coded by physicians and classified according to the ICD-8 until the end of 1993 and the ICD-10 thereafter…. We excluded patients who were younger than 18 years old on the hospital admission date (the index date) and who had a hospital contact with acute pancreatitis between 1977 and 2004.

Population Control Subjects—We used the Civil Registration System to select 10 population control subjects for each case. Control subjects were individually matched on birth year, sex, index date, and Danish region of residence. We selected control subjects using risk-set sampling from persons alive and at risk for a first hospitalization with acute pancreatitis on the case patient’s index date.”
“Use of Incretins and Other Antihyperglycemic Drugs—The Danish National Drug Prescription Registry contains information on reimbursed prescriptions dispensed at all pharmacies in Denmark, including date of sale, active substance, route of administration, and amount dispensed. We retrieved data from the NDRP from 2004 until the index date for both case patients and control subjects. Ever use of incretins was defined by one or more prescriptions for DPP4 inhibitors or GLP-1 receptor agonists as monotherapy or combination therapy, regardless of other comedication.

Current use of incretins or other antihyperglycemic drugs was defined by at least one relevant prescription within 100 days before the index date. Former use was defined as redeeming no prescriptions within 100 days and at least one prescription >100 days before the index date. Nonuse was defined as no recorded prescription at any time before the index date.”

Case-Crossover Design (A Case-Only Design)
Pre-Post Design

Target population

Study population

Unexposed (time)

Exposed (time)

observer

time

outcome
Interrupted Time Series

Target population

Study population

Unexposed
(time)

Exposed
(time)

Outcome rate
(interval)

Observer
Rigorous analyses begin with:

- Definition of study population
- Definition of exposures
- Definition of outcomes