LINKING EXISTING DATABASES: POISONED CHALICE OR HOLY GRAIL?

Dan Huse
Thomson Reuters

ISPOR 14th Annual European Congress
Madrid, Spain
November 7, 2011

BACKGROUND:
U.S. HEALTH CARE DATA

• Health care delivery in the US is mainly reimbursed on a fee-for-service basis
• Each visit, prescription, procedure, test, etc., generates a transaction record (paid claim)
• Large claims databases are widely available for research purposes
• HOWEVER, limited clinical information is required to justify payment
• HENCE, linkage to more detailed clinical data is valuable

DISCLOSURES

• Dan Huse is an employee of Thomson Reuters. The company’s MarketScan Research Databases® are discussed in this presentation.

EXAMPLE:
MARKETSCAN RESEARCH DATABASES

• Claims data from US employers, health plans, and state governments have been linked to:
  – Laboratory test results
    • Electronic data capture from national laboratories
  – Health risk assessments
    • Self-reported data from workplace wellness programs
  – Hospital inpatient records
    • Discharge summaries and detailed service utilization
  – Dates of death
    • From Social Security Administration files
  – Office-based medical records
    • EMR data from general medical practices

LINKING TO DEATH RECORDS:
HOLY GRAIL?

• Benefits:
  – Mortality is an important outcome of many chronic diseases
  – Death dates will allow us to distinguish complete longitudinal records from censored ones

• Feasibility:
  – Social Security Administration provides records of deaths of persons registered with Social Security
  – Employers represented in MarketScan commonly use Social Security identifiers in health plan records

LINKING TO DEATH RECORDS:
POISONED CHALICE?

• Not all patients have the identifying information to use for linking
  – Exclude records from employers not supplying adequate personal identifiers
  – Exclude records from employers supplying personal identifiers of uncertain accuracy
  – Exclude spouses and children due to data quality concerns
  – Require corroborating evidence, such as name or date of birth

• Observed versus expected death rates varied over time
  – Exclude older data from linkage

OBSERVED VERSUS EXPECTED MORTALITY

LINKING TO OFFICE-BASED EMR DATA:
HOLY GRAIL?

• Benefits:
  – EMR has clinical information not found in claims
    • Vital signs (e.g. BP, heart rate, temp)
    • Weight, height, BMI
    • Habits (e.g. smoking, alcohol use)
    • Other biometrics (EKG, BMD, FEV, etc.)
    • Laboratory results
    • Prescriptions as written

• Feasibility:
  – Partnership with leading EMR supplier, General Electric
  – Significant geographic overlap between GE and TR data capture
LINKING TO OFFICE-BASED EMR DATA: POISONED CHALICE?

- US privacy laws are barrier to sharing personal identifiers between business partners
  - Also business agreements prevented sharing of physician IDs
- Linkage based on:
  - Month and year of birth
  - Sex
  - ZIP3 (postal code)
  - Dates of physician visits
- Monte Carlo simulations used to estimate and minimize risk of false positive matches
  - Minimum of 3 matching visit dates
- Result is convenience sample biased toward frequent users of medical care

POISONED CHALICE OR HOLY GRAIL?

- Limitations:
  - Privacy laws and other governance issues often preclude use of personal and provider identifiers
  - Availability of identifiers does not guarantee linkage if data recording is of variable quality
  - Scope of linkage is therefore constrained and subject to selection bias
- Value of linked patient records
  - Rich clinical AND economic data source
  - Transparent linkage methodology and results enables analysts to use data appropriately
  - Over time, linkage methods improve and sample sizes grow

CONTACT INFORMATION

- Dan.Huse@thomsonreuters.com