Use and Management of Big Data in Health Economics and Outcomes Research in Korea

SeungJin Bae
Associate Professor
Ewha Womans University, College of Pharmacy

Overview

Brief description of big data in Korea
• Detailed characteristics of NHI-related data
  – HIRA data
  – NHIS
• Application of the big data in HEOR
Big data in Korea

• Available upon request (sometimes review required)
  – Korean NHI Claim data
    • HIRA/NHIS
  – Cause of Death data
    • Department of statistics
  – Cancer registry
    • National cancer center

General description of Korean NHI data

• Health Insurance claims data
  – 97% of total population
• Medical assistance program
  – 3% of total population
• Electronic Claims over 90% since 2005
  – Almost 99% as of 2009

• General Information
  – Date of birth, gender
  – Patients’ cost sharing, insurer’s payment amount

• FFS payment
  – Medical services/devices/medicine’s itemized
  – Rx detail
### Pros and cons

#### Pros
- Represent total population
  - Burden of disease estimation
  - Health care utilization pattern
- Comprehensive coverage
- Represent current practice pattern

#### Cons
- Only Captures utilizations covered by the NHI
  - Non-reimbursed services not measured
  - 28% of healthcare expenditure in non-reimbursed sector
  - Patients who has the disease but do not seek care could not be captured
  - Could underestimate the burden of diseases (e.g., migraine)
- Internal Validity issue
  - About 70% of primary diagnosis concurred with Medical Record
  - Consistency of the clinical definition
    - Subjective diagnosis criteria
    - GERD vs. esophagitis/dyspepsia

---

### Korean NHI: HIRA data

#### Data use
- Approval required
  - On-site use: big data (more than sample)
  - Remote access (started since 2013)
- Not required
  - Sampled cohort: CD distribution
NHIS–NSC cohort

- National Health insurance service–national sample cohort
- A population–based cohort
- Constructed to provide researchers & policymakers with representative data
  - Healthcare utilization, health examination, income
- 2.2% of total population as of 2002 (1,025,340)
  - Stratified sampling (of 1,476)
- Followed for 11 years (until 2013)
Application in HEOR

- Identify target population
  - Diagnosis code (ICD codes)
  - Disease-specific procedures/drugs

- Incidence rate, health care utilization, health outcome estimation (Bae et al., 2012)
Application in HEOR (2)

• GERD prevalence study
• Definition of GERD:
  - primary or secondary disease
the Korean Standard Class
• Healthcare utilization
  - Proton pump inhibitor

Thank you

• References
