Use of Automated Databases for Pharmacoepidemiology Research

Alison Bourke BSc MSc FRPharmS FISPE
Scientific Director, Center for Advanced Evidence Generation, IQVIA Real-World & Analytics Solutions

6th MURIA – July 8, 2019

Acknowledgements & Conflict of Interest

Thanks to:-
Vincent Lo Re, MD, MSCE, FISPE – Adapted slides from Muria 2018
Perelman School of Medicine, University of Pennsylvania

Conflict of Interest:-
I am a part time employee of IQVIA within the Real World & Analytics Solutions group

Learning Objectives

- Review databases for pharmacoepidemiology research
  - Registries, claims, medical records
  - Understand their strengths, weaknesses
- Facilitate appropriate database selection
  - Clarify reasons for database selection

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology:
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology:
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Over 30 years of Databases

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

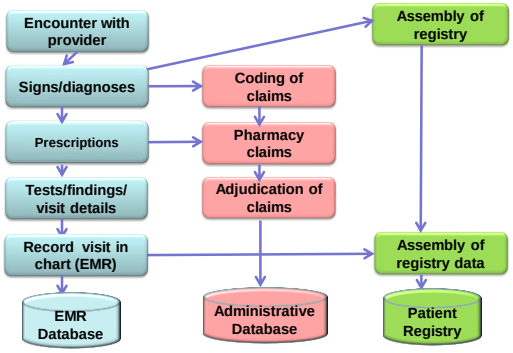
Automated Databases

- Allow evaluation of health conditions in “real world” settings
- Past three decades → ↑ use of electronic data sources containing medical care data
- Efficient, cost-effective way to conduct pharmacoepi research

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Automated Database Generation



Schneeweiss S. *J Clin Epidemiol* 2005;58:323-37

Ideal Automated Database

Ideal Automated Database 1

- Longitudinal data from all care settings
- Records prescribed, dispensed drugs
- Includes laboratory tests results
- Confounders of interest available
- Validated data
- Ability to collect additional information from care team and patients

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Shah BR. *Am Heart J* 2010;160:8-15.

Ideal Automated Database 2

- Linkable to other data sources (via identifiers)
- Large representative population
- Access to experts in data context
- Stable
- Updatable, with access to medical records

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Shah BR. *Am Heart J* 2010;160:8-15.

"Database studies must be performed within the limitations of a resource not specifically designed to test the research hypothesis"

Gillian C. Hall, PhD

Choosing Among Databases

Key Point: The research question dictates selection of the appropriate pharmacoepidemiology database.

- Appropriate study population, size
- Ascertain exposure, outcome
- Relevant confounders measured
- Link with other databases, records

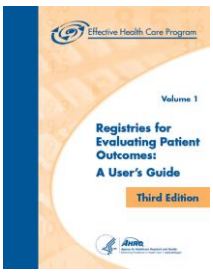
Pharmacoepidemiol Drug Saf. 2012 Jan;21(1):1-10. doi:10.1002/pds.2229. Epub 2011 Nov 8

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

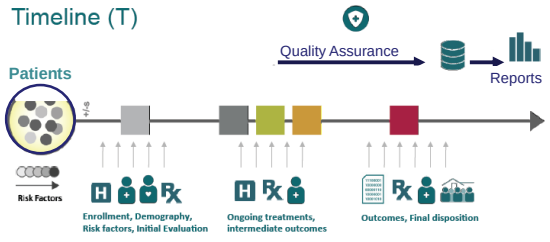
Registry Databases: Overview

- Prospective study of patients with common characteristics
- Developed to evaluate:
  - Natural history of disease
  - Drug effectiveness, safety
  - Quality of life
  - Cost-effectiveness of therapies



Gliklich RE, Dreyer NA. AHRQ publication No. 07-EHC001.
http://effectivehealthcare.ahrq.gov/ehc/products/21/12/PatOutExecSumm.pdf

Development and Maintenance of Patient Registries



http://www.pcori.org/assets/11-Gliklich-Slides-Registries.pdf

Registry Databases: Data Collected

- Collect data on:
  - Demographic characteristics
  - Social history
  - Disease-specific drug treatments
  - Select disease-related outcomes
- Ability to link to other data sources?

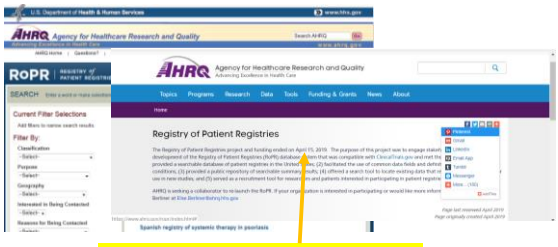
Registry Databases: Benefits

- Large patient numbers
- Usual diagnostic, follow-up procedures
- Contain "Real World" therapeutic effectiveness, safety data
- Heterogeneity among sites

Registry Databases: Limitations

- Selection bias (non-sequential patients)
- Variability in data definitions
- Data may not be validated
- Incomplete data on comorbid conditions, outcomes, mortality
- Inability to link with other data sources

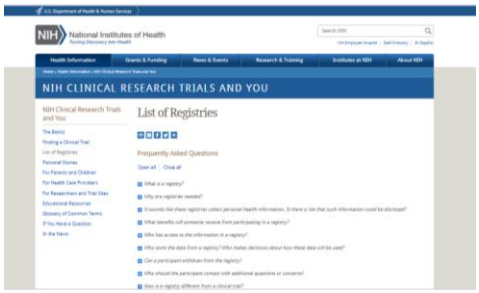
Where to Find Registries?



The Registry of Patient Registries project and funding ended on April 15, 2019

https://patientregistry.ahrq.gov/

Where to Find Registries? USA?



https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries

Where to Find Registries? Europe?

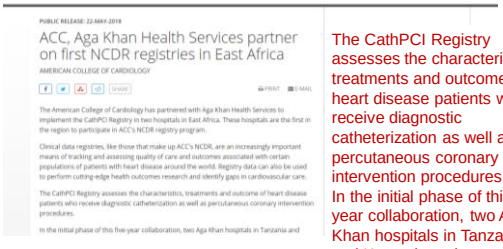
https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries

Where to Find Registries? Asia?

ADORE is a real-world, prospective, longitudinal, investigator-led registry that will collect comprehensive data for type 2 diabetes treatment patterns and outcomes throughout Asia with a goal to create a network of sites with thousands of patients in China, South Korea, Taiwan, Singapore, Indonesia, Malaysia, Thailand, the Philippines, Hong Kong, India, Vietnam and Sri Lanka.

https://www.acc.org/about-acc/press-releases/2015/11/05/10/04/us-and-asian-registries-combine-to-create-first-global-diabetes-registry

Where to Find Registries? Africa?



PUBLIC RELEASE: 22-MAY-2018

ACC, Aga Khan Health Services partner on first NCDR registries in East Africa

AMERICAN COLLEGE OF CARDIOLOGY

The American College of Cardiology has partnered with Aga Khan Health Services to implement the CathPCI Registry in two hospitals in East Africa. These hospitals are the first in the region to participate in ACC's NCDR registry program.

Clinical data registries, like those that make up ACC's NCDR, are an increasingly important means of tracking and assessing quality of care and outcomes associated with certain populations of patients with heart disease around the world. Registry data can also be used to perform cutting-edge health outcomes research and identify gaps in cardiovascular care.

The CathPCI Registry assesses the characteristics, treatments and outcome of heart disease patients who receive diagnostic catheterization as well as percutaneous coronary intervention procedures.

In the initial phase of this five-year collaboration, two Aga Khan hospitals in Tanzania and

The CathPCI Registry assesses the characteristics, treatments and outcome of heart disease patients who receive diagnostic catheterization as well as percutaneous coronary intervention procedures.
In the initial phase of this five-year collaboration, two Aga Khan hospitals in Tanzania and Kenya have been selected to participate in the CathPCI Registry.

https://www.eurekalert.org/pub_releases/2018-05/acoc-aak052218.php

Where to Find Registries? Africa?

Cancer Registries

The Kampala registry in Uganda
- 3 staff: 1 Director and 2 full time staff – registry manager and data clerk
- Population of 2 million (about 1,000 new registrations per year)

The Harare registry in Zimbabwe
- 5 staff: 1 Registry manager, 1 executive assistant and 3 health information assistants
- Population of ca. 1.5 million and passive coverage of the whole country (approximately 5,000 new cases a year).

http://www.eurekalert.org/pub_releases/2018-05/acoc-aak052218.php

How to build Registries

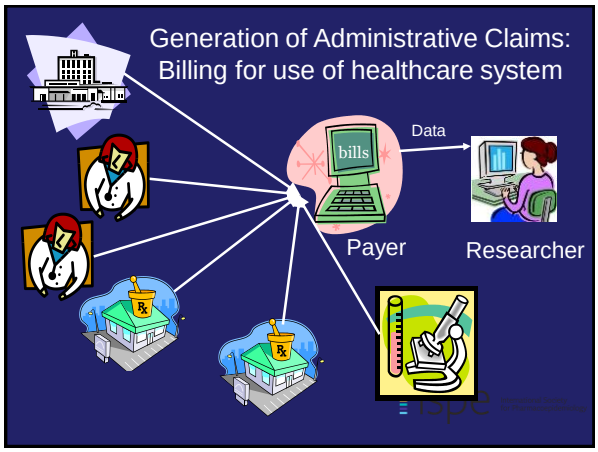
UICC Cancer Registries- why what how.pdf

file:///C:/Users/abourke/Documents/Important%20Documents/AB%20OfficeStuff/Presentations/2019/UICC%20Cancer%20Registries-%20why%20what%20how.pdf

http://www.eurekalert.org/pub_releases/2018-05/acoc-aak052218.php

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection



Examples of Claims Databases

- US government: US Medicaid, Medicare
- US commercial insurance
- Canadian provincial
- Italian Agencia Regionale di Sanita Tuscany
- French Securite Sociale de l'Assurance Maladie
- German Pharmacoepidemiological Research Database
- Japan Medical Data Centre Claims database

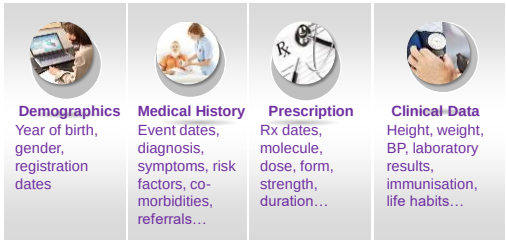
Advantages & Disadvantages of Claims Databases

Advantages	Disadvantages
• Holistic care	• Not recent
• Accurate…coders?	• Condition limited
• Includes dispensed Rx	• Rx formulary
• Diagnosis-Rx link	• Selected population
• Included costs	• Reduced data items

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

Patient Data Collected



Demographics
Year of birth, gender, registration dates

Medical History
Event dates, diagnosis, symptoms, risk factors, co-morbidities, referrals…

Prescription
Rx dates, molecule, dose, form, strength, duration…

Clinical Data
Height, weight, BP, laboratory results, immunisation, life habits…

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Where to Find EMR Databases?

ispe International Society for Pharmacoepidemiology

ww.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries

Examples: Clinical Practice Research Datalink (CPRD) & The Health Improvement Network (THIN)

- United Kingdom medical record databases
- 30+ years experience, 1000s publications
- General practitioner: “gatekeeper”
- Available data:
  - Medical diagnoses
  - Outpatient prescriptions
  - Lab results
- Hospital care → Linked Hospital Episode Statistics

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

http://www.cprd.com/

EMR Databases: Strengths Population

- Relevant clinical data
- Large, real-world, representative population
- Includes vulnerable (eg elderly, children, pregnant)
- Population based
- Longitudinal
- Historic data
- No recall bias
- Wide formulary

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

EMR Databases: Strengths Operationally

- Usually pre coded data
- Frequently updated
- Links to other info
- Study design flexibility
- Long follow up
- Fast – Information already collected
- Less Expensive

Potential Limitations of EMR Databases 1

- Uncertain validity of diagnoses
- Completeness, quality of data
- Instability of population
- Generalizability
- Costs of data
- Unlikely to cover all healthcare settings
- Geographic & care diversity
- Confidentiality and privacy – vital, governance may be unclear

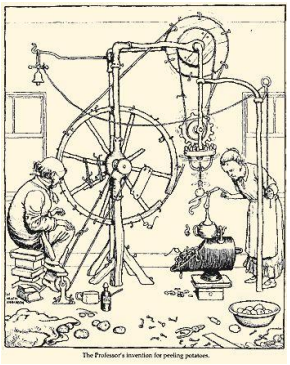
Potential Limitations of EMR Databases 2

- Need expertise in context of data collection
- Need robust analytics & IT
- Generally no direct labelling of indication
- Limited information on OTC medications
- Limited data on non-routine care, lifestyles, diet
- Limited information on how patients feel, what is important to them
- May not be dispensed prescriptions

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

Hybrid/Enriched Databases



Heath Robinson
Potato Peeler

ispe International Society for Pharmacoepidemiology

Outline

- Overview of automated databases
- Data sources for pharmacoepidemiology
  - Registries
  - Claims databases
  - Electronic medical record (EMR) databases
  - Hybrid/Enriched databases
- Appropriate database selection

IMS Health & Quintiles are now IQVIA

ispe International Society for Pharmacoepidemiology

Choosing Among Databases

- Research question dictates database
- Existing "checklists" to guide researchers
  - ISPE guidelines
  - ISPOR guidelines

Hall GC. *Pharmacoepidemiol Drug Saf* 2012;21:1-10.
Berger ML. *Value Health* 2009;12:1053-61.

PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2012; **21**: 1–10
Published online 8 November 2011 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.2229

COMMENTARY

Guidelines for Good Database Selection and use in Pharmacoepidemiology Research†

Gillian C. Hall[1]*, Brian Sauer[2], Alison Bourke[3], Jeffrey S. Brown[4], Matthew W. Reynolds[5] and Robert Lo Casale[6]

[1] *Grimsdyke House, London, UK*
[2] *Salt Lake City VA IDEAS Centre & Division of Epidemiology, The University of Utah, Salt Lake City, UT, USA*
[3] *CSD Medical Research, London, UK*
[4] *Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA*
[5] *United BioSource Corporation, Lexington, MA, USA*
[6] *Department of Epidemiology, Merck & Co. Inc, West Point, PA, USA*

International Journal of Epidemiology
© International Epidemiological Association 1996

Vol. 25, No. 2
Printed in Great Britain

A Framework for Evaluation of Secondary Data Sources for Epidemiological Research

HENRIK TOFT SØRENSEN,* SVEND SABROE** AND JØRN OLSEN†

Sørensen H T (Department of Internal Medicine V, Aarhus University Hospital DK-8000 Aarhus C, Denmark) Sabroe S and Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *International Journal of Epidemiology* 1996; **25**: 435–442.

Important Questions to Ask

- What is the population covered?
- Are there continuous, consistent data?
  - Exposure, outcomes
  - Confounders of interest
- Is follow-up sufficiently long enough?
- Access to medical records?
- Ability to link to other data sources?

4 Key learning to take away

- All databases have strengths, limitations
- Research question guides database selection
- Understand accuracy, completeness, appropriateness of data
- Collaborate with experts in data sources

Any questions?
Your experience?

Thank you

alison.bourke@iqvia.com