

OBJECTIVE

To provide an example of how the online resource B.R.I.D.G.E. TO DATASM serves as an education tool in (pharmaco)epidemiologic and pharmaco-economic research, and to understand and identify these data sources worldwide.

BACKGROUND

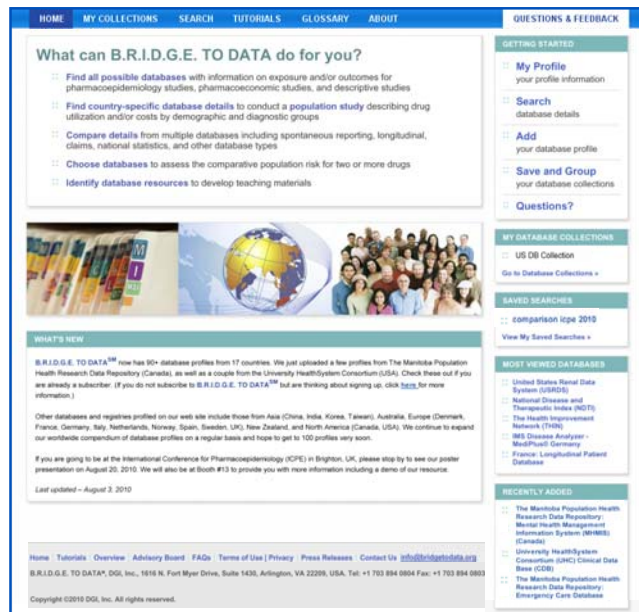
Teaching pharmacoepidemiology research requires tools to instruct students on:

1. Identifying and accessing multiple databases for evaluating optimal data resources to address a research question, and
2. Facilitating understanding of the structure and data elements in healthcare databases.

One such resource is B.R.I.D.G.E. TO DATASM (Benefit-Risk Information for Drug Evaluations, (www.bridgetodata.org; Figure 1), a restructured online database of databases originally known as the Risk Assessment of Drugs Analysis & Response (RAD-AR) Handbooks.¹

¹ Miwa LJ, Jones JK, Mann RD. Identifying resources for assessing use and effects of pharmaceuticals and other medical products in the European community. *European Medicines Research*. IOS Press; 1994. p. 391-407.

Figure 1. B.R.I.D.G.E. TO DATASM Homepage



METHODS

Where are the data? A researcher may hypothesize that there is a difference in clinical outcome (i.e., drug-induced respiratory diseases) when a certain drug is administered at two different dosage levels (high; low). To do this study in the US, the researcher would need to identify US databases that capture data on drugs (with dosages) and diagnoses. Table 1 lists data elements contained in the database profiles in B.R.I.D.G.E. TO DATASM.

Table 1. Brief Description of Data Elements

Category	Description
Summary	Database history, uses, strengths / limitations
Population Dynamics	Population size (active/final), annual change in size, sample weights
Demographic Data	Age, gender, ethnicity/race, geographic location, date of birth, death data
Physician & Practitioner Info	Physician ID & specialty; Pharmacy ID
Diagnoses/Signs & Symptoms	Date parameters, coding system(s), as well as data on physical exams, birth defects, cancer, infectious diseases, environmental exposures, and behavior
Procedures	Date parameters, coding system(s), lab data
Drug Information	Date parameters, regimen & route, manufacturer, dosage, coding system(s), generic name
Economic Data	Cost data including type; surrogate cost data
Validation & Linkage	Database linkage capabilities to and validation against other sources; access to medical records
Administrative Data	Contact and accessibility information for each database, recent references, and date the database was last updated

Two types of searches were simulated in bridgetodata.org to identify databases that contain appropriate diagnosis & drug data and are large enough to study respiratory disease outcomes. First, a criteria-based search was conducted to identify databases by categories of information (Figure 2). To supplement this, a keyword-based search was also conducted.

CONCLUSION

This example demonstrates that bridgetodata.org supports decision-making and identification of a large number of databases worldwide for (pharmaco)epidemiologic and pharmaco-economic research. It also serves as a teaching tool for understanding healthcare databases and their attributes.

RESULTS

Figure 2. Criteria-based search to find optimal US database(s) for studying the risk of developing respiratory diseases in patients exposed to medication at varying dosages.

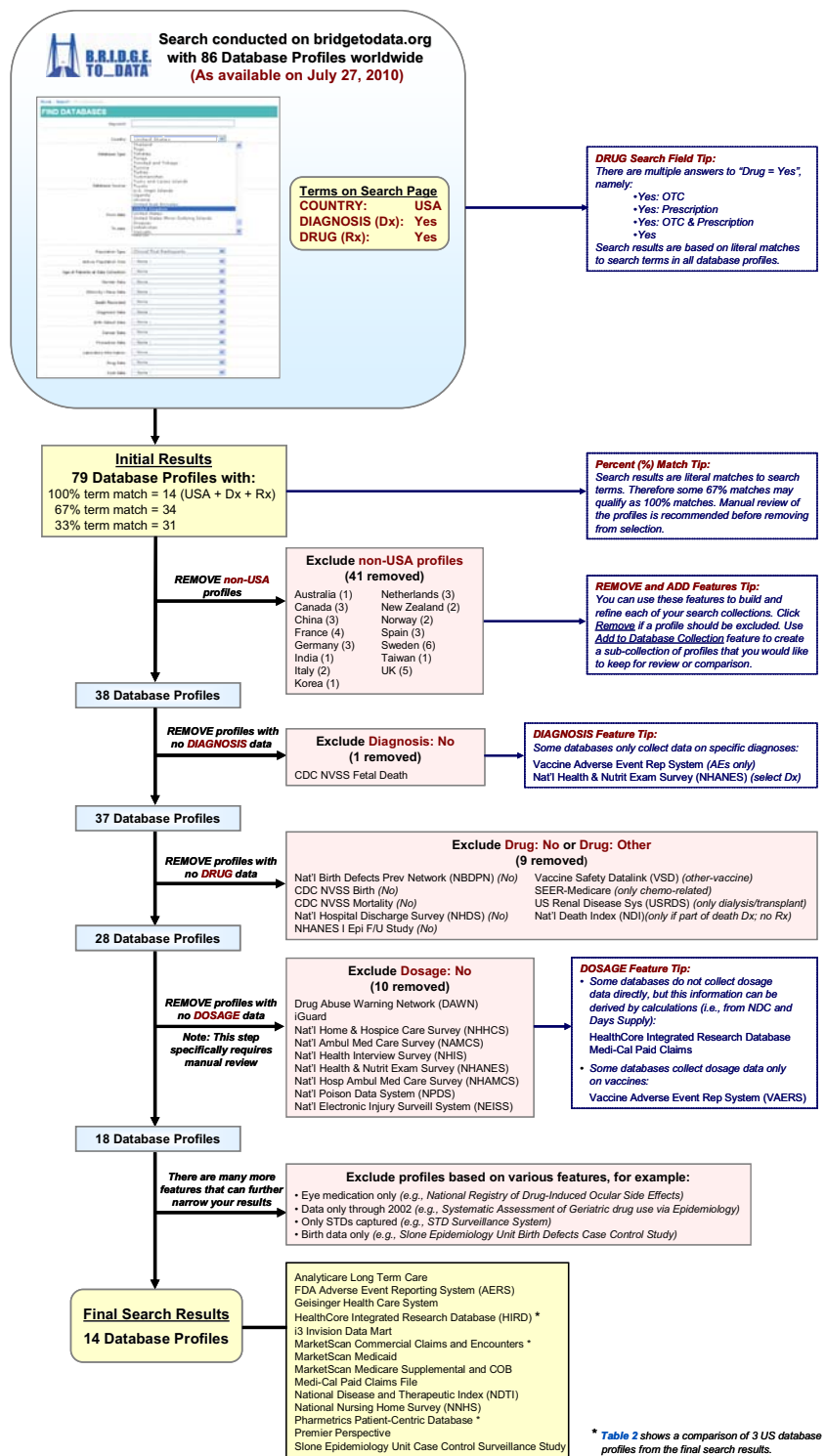


Table 2. Excerpt from B.R.I.D.G.E. TO DATASM comparing 3 of the selected US databases

FIELD NAMES	HEALTHCORE INTEGRATED RESEARCH DATABASE (HIRD)	MARKETSCAN COMMERCIAL CLAIMS AND ENCOUNTERS	PHARMETRICS PATIENT-CENTRIC DATABASE
Demographic Data			
Population Type	Insured: type (e.g., Medicare, Medicaid) Private employer-based insurance	Insured: type (e.g., Medicare, Medicaid) Active employees and dependents, early (non-Medicare) retirees and dependents, and COBRA continuers	General Population Commercially insured population
Approximate Percentage of Participants <18 years and those >65 years	< 18 = 20% > 65 = 11%	N/A	<18 = 25.5% >65 = 65.6%
Percentage of Males/Females	Males = 49% Females = 51%	N/A	Males = 50% Females = 50% (80% females and 40% males have received services)
Geographic Location	Southeast Mid-Atlantic Eastern Central Western regions (Only those in U.S. health plans)	All regions of USA (The database contains information at region, state, and 3-digit zip code level)	U.S. Census Regions - East, Southeast, Midwest, West
Date of Birth Recorded	Yes (Date / Month / Year)	Yes (Year)	Yes (Year)
Death Recorded	No Only date of death is available by linking to an external source	Yes Death information is available on hospital discharge records	No
Other Demographic Data	No	Yes Information such as relationship to policy holder, employment status and family history (cancers and other chronic conditions) data can be obtained from the Health Risk Assessment Database	No
Population Dynamics			
Database Population Size	20 - 50 Million (39 million - number of people who have existed in the database at any time - Fully Insured and ASO. For just "Fully Insured" it would be 31.4 million.)	50 - 100 Million (73.9 million)	50 - 100 Million (~60 million patients)
Active Population Size	5 - 20 Million (12.9 million ASO & Fully Insured; 9.5 million Fully Insured only)	N/A	5 - 20 Million (~16 million)
Diagnoses/Signs & Symptoms			
Diagnosis Data	Yes	Yes	Yes
Diagnoses Coded	ICD-9-CM	ICD-9-CM DRG	ICD-9-CM
Diagnoses: Maximum Number of Codes Allowed	5	2 (2 diagnoses per inpatient or outpatient procedure)	4 (4 per record)
Procedures			
Laboratory Information	Yes	Yes Biometric data such as BMI, blood pressure, cholesterol, etc. are available. Outpatient laboratory results are available via linkage to the Lab Database.	Yes Lab results are available for a subset of 1 million patients
Drug Information			
Drug Data	Yes: Prescription only	Yes: Prescription only Filled prescriptions (retail or mail order), physician administered (specialty, immunizations/vaccines, etc.)	Yes: Prescription only
Drug Coding System: Primary	NDC	NDC	NDC Mapping tables to AHFS, USC available
Drug Dosage	No (Only through calculation of NDC (mg specific) and days supply)	Yes	Yes NDC gives dose form, but not number prescribed or Slg

Educating a Pharmacoepidemiology Researcher on Selecting Optimal Healthcare Databases for Research

Sharmila A Kamani, Earl L Goehring Jr, Anokhi J Kapasi, Bao Nguyen-Khoa, Varinder P Singh, Judith K Jones

DGI, Inc., Arlington, VA, USA

Presented at the 26th International Conference on Pharmacoepidemiology & Therapeutic Risk Management.
August 19-22, 2010, Brighton, UK.



DGI Center for Health Research and Education Inc.

1616 N Fort Myer Drive, Suite 1430 • Arlington, VA 22209 USA

www.bridgetodata.org

T: +1 703 894 0804

info@bridgetodata.org