Comparing and Identifying Optimal Healthcare Databases for Comparative-Effectiveness of Breast Cancer Screening

A.J. Kapasi, **S.A. Kamani,** B. Nguyen-Khoa, V.P. Singh, J. Tao, E.L. Goehring Jr, **J.K. Jones**
DGI, LLC, Arlington, VA, USA E-mail: info@bridgetodata.org

BACKGROUND

- Comparative effectiveness research (CER) involves a thorough understanding of optimal resources for evaluating and comparing health outcomes and clinical effectiveness of medical treatments or health services.
- This requires familiarity with epidemiology, patient populations in various settings, as well as treatment patterns.
- A powerful step in achieving quality CER is to know which healthcare databases are available and their strengths and limitations.
- B.R.I.D.G.E. TO DATA® (B.R.I.D.G.E.; www.bridgetodata.org) an international resource of database profiles, may serve as one resource for CER studies.

OBJECTIVE

To show how researchers may identify and compare multiple healthcare databases by using a CER-based question of high public health impact.

METHODS

Case Study: CER analysts must determine whether differences exist in breast cancer diagnoses, health outcomes, and costs between women (<50 years) who receive mammograms, and those who do not.

A search was conducted in www.bridgetodata.org to identify databases collecting necessary data as shown in **Figure 1**.

Figure 1. B.R.I.D.G.E. TO DATA® Search Page

Age = Yes
Gender = Yes
Diagnosis Data = Yes
Cancer = Yes
Laboratory Data = Yes
Cost Data = Yes

Keyword:
Country: Please select
Database Type: - None -
Database Source: - None -
To search by Years Covered, please enter a date range. The "To date" is optional, but the "From date" is required to perform date searches.
From date:
Format: 2012
To date:
Format: 2012
Population Type: - None -
Active Population Size: - None -
Age of Patients at Data Collection: Yes
Gender Data: Yes
Ethnicity / Race Data: - None -
Death Recorded: - None -
Diagnosis Data: Yes
Birth Defect Data: - None -
Cancer Data: Yes
Procedure Data: - None -
Laboratory Information: Yes
Drug Data: - None -
Cost Data: Yes
Data Validation Against Original Source: - None -
Access to Medical Records: - None -
Filter Results Reset

A 100% match was identified either by a (i) 100% relevancy ranking OR (ii) adjudication of profiles with an 83% match (5/6 criteria) using supplemental information.

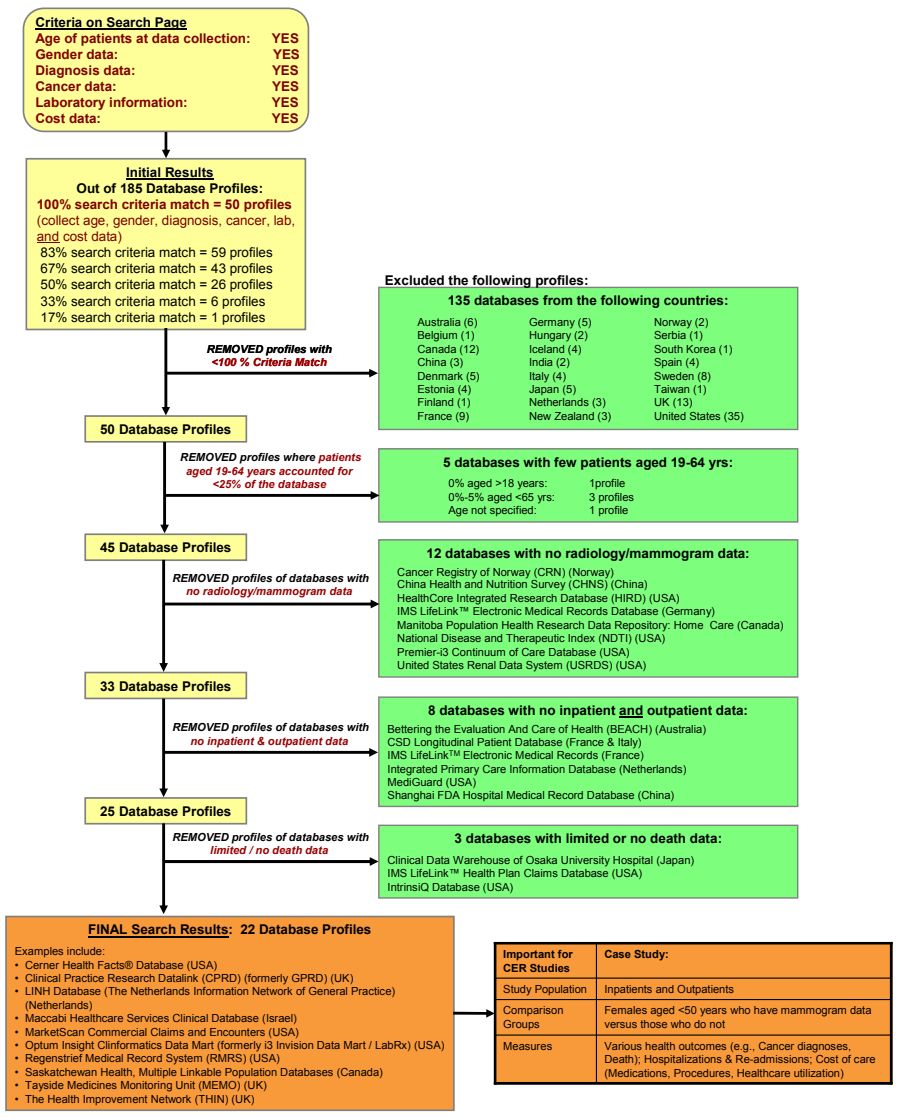
Search results were further narrowed by excluding databases with no data on adults, mammograms, hospitalization, or death. Databases that have linkage capabilities to obtain relevant data were included.

RESULTS – Part 1

- The search comprised of 185 databases from 25 countries (including 27 from Asia, Middle East, Australia, and New Zealand) and 14 oncology-specific databases.
- The initial search yielded: 50 databases matching 100% of the search criteria;135 databases not matching all 6 criteria were excluded (**Figure 2**).
- Exclusions from the remaining 50 databases were made if:
  - Patients **aged 19-64 years** accounted for <25% of the database population (n=5)
  - **Radiology** or **mammogram** data were lacking (n=12)
  - Both **inpatient** and **outpatient** data were not included (n=8)
  - **Death** information was unobtainable (n=3).
- Of the final set of 22 databases, the most common cost data included those related to medication (n=12) and/or hospital (n=14) utilization, although the type of cost information varied (e.g., billing, co-pay, reimbursement).

RESULTS – Part 2

Figure 2. Criteria-based search conducted in www.bridgetodata.org for CER case study
(185 Database Profiles worldwide as of October 4, 2012)



These profiles can be compared side-by-side to identify the most appropriate database(s) for answering the proposed CER question (**Table 1**). Additional data for consideration may include collection of ethnicity/race data, data access, and validation of data.

Table 1. Excerpt from B.R.I.D.G.E. comparing data elements in 3 selected databases

FIELD NAMES	Japanese Breast Cancer Registry (Japan)	National Central Cancer Registries (NCCR) (P.R.China)	New Zealand Cancer Registry (NZCR)
Brief Database Description	The Japanese Breast Cancer Registry has been developed and is maintained by the Japanese Breast Cancer Society since 1975. A total of 188,265 breast cancer patients were registered 1975-2003; after introducing a new registration system in 2004, another 207,468 cases were registered 2004 - February 2012. The Registry thus has a total of 395,773 cases since 1975. The data are collected from 851 nationwide facilities. The registry is considered to cover 70% of patients with breast cancer in Japan. [more...]	The National Central Cancer Registries in the People's Republic of China contain data from 48 cancer registration areas that are distributed in 19 provinces (autonomous regions and municipalities), comprising 20 cities and 28 counties. Cancer registration in urban areas largely covers central districts but not suburbs, except for Nangang District and Daoli District covered in Harbin Cancer Registry, and all districts besides Tianhe District in Guangzhou registry. Cancer registries in rural areas cover all areas in the county. Cancer clinical and personal information of patients with cancer are collected with registry's population data stratified by age group and sex in mid-year. The Registry contains data since 1988 and currently has data on about 71 million cancer patients. [more...]	The New Zealand Cancer Registry (NZCR), set up in 1948, is a population-based register of all primary malignant tumours diagnosed in New Zealand, excluding squamous and basal cell skin cancers. The Cancer Registry Act 1993 and Cancer Registry Regulations 1994 were introduced to increase reporting of primary cancers in New Zealand; primary information comes from pathology reports. Since this legislation, laboratory test results have been collected, and data quality and completeness have significantly improved. There are about 18,500 new registrations of cancer each year, with the highest rates in the middle and older age groups. [more...]
Database Type	Specific Disease Registry (Breast Cancer Registry)	Specific Disease Registry (NCCR covers 48 registration areas in the country)	Specific Disease Registry (Cancer Registry)
Database Source	Medical Records (Data are collected from 851 nationwide facilities in Japan)	Medical Insurance Claims Medical Records Other sources include: Vital statistics from Civil Administration Bureau	Primarily from pathology reports and public hospitals. Additional data sources include: Medical Certificates of Cause of Death, Coroners' Findings/autopsy reports, hospital discharge data on the National MDS, private hospital discharge returns.
Years Covered	1975 - Present	1988 - Present	1948 - Present
Patient Type	Inpatient and Outpatient	Inpatient and Outpatient	Inpatient and Outpatient
Database Population Size (Range)	<0.5 Million (395,773 breast cancer patients)	50 - 100 Million (~71 Million)	1 - 5 Million
Active Population Size	<200,000 (Bilateral cancer counts as 2 cases in this registry)	~5 million (in the past 10 years) [more...]	<0.5 Million
Approximate Percentage of Participants <18 years and those >65 years	<18 years = 0.0% >65 years = 31.8% (Per the 2008 Annual Report)	< 20 years = 65% (12.3 Million) > 65 years = 35% (6.6 Million) (Based on the 2010 Annual Report, which has data from 2007.)	N/A (Age-specific rates for cancer registrations were considerably higher in older age groups in 2005)
Percentage of Males/Females	Females = 95.5% Males = 0.5% (Per the 2008 Annual Report)	Males = 50.75% (~30.5 Million) Females = 49.25% (~29.6 Million) (Based on the 2010 Annual Report, which has 2007 data from 38 registries)	Males = 52% (9,647 as of 2005) Females = 48% (8,963 as of 2005) (The rates have lowered since 2004 and are significantly less than in 1995) [more...]
Death Recorded	Yes: date of death is recorded	Yes	Yes: In 2005, New Zealand recorded 7,971 cancer deaths. Deaths from cancer have decreased by 2.1% from 2004 and increased by 7.4% from 1995. Cancer remained the leading cause of death, accounting for 29.4% of deaths, slightly more than non-congenital heart diseases.
Diagnoses Coded	ICD-10 Information includes the way the cancer was detected (self-detection, medical exam), as well as stages, and tumor size.	ICD-10 ICD-O-3	ICD-10 (used to code tumour site) ICD-O (used to code tumour morphology) ICD-9
Cancer Data	Yes Information includes the way the cancer was detected (self-detection, medical exam), as well as stages, and tumor size	Yes Information is included on cancer sites as follows: Lip, Tongue, Mouth, Salivary Glands, Tonsil, Other Oropharynx, Nasopharynx, Hypopharynx, Pharynx Unspecified, Oesophagus, Stomach, Small Intestine, Colon, Rectum, Anus, Liver, Gallbladder etc., [more...] No cancer stage information is available as of now, but data collection for this information is in process in some provinces.	Yes NZCR collects as much information as possible about the tumour - most of this information comes from pathology reports and includes data items such as prognostic markers for breast cancer (e.g., her2 status). Staging information is collected where available. Since pathology reports are the main source of data, this information is most complete for tumours where the primary treatment is surgical (e.g., melanoma, breast, colorectal cancers).
Procedure Data	Yes Information is captured on surgical procedures	Yes	No Only collects cancer incidence data, no treament data. [more...]
Laboratory Information	Yes This includes pathology of the tumor and lymph node metastases	Yes	Yes Laboratories are the primary source of cancer data for NZCR. Any new diagnosis of cancer in New Zealand must be reported, excluding squamous and basal cell skin cancers. Tumour marker data are also collected where provided on the pathology report, for example, Estrogen, Progestin, HER2 +/-). [more...]
Drug Data	Yes: Prescription only	No	No
Cost Data	No	No	No
Type of Cost Data	N/A	N/A	N/A
Brief Description of Linkage Capabilities	Direct linkage exists to the National Clinical Database	NCCR is linked to the Vital Statistics Database	NZ has a unique personal health identifier - National Health Index (NHI) - which allows linkage to several databases held at the Ministry of Health, including mortality. [more...]
Database Contact Data	The Japanese Breast Cancer Society [Full contact information in B.R.I.D.G.E.]	Dr. Wanqing Chen [Full contact information in B.R.I.D.G.E.]	Ms. Susan Hanna [Full contact information in B.R.I.D.G.E.]
Source of Database Funding	Private (The Japanese Breast Cancer Society)	Government (National Central Cancer Registry)	Government (New Zealand Ministry of Health)
Database Usage Restrictions	Private Access (Database access is restricted)	Private Access (Data cannot be accessed by external researchers; however, they may apply for data reports and work in collaboration with the database reseachers.)	Public & Private Access: The NZCR system is only accessed by authorised Ministry of Health staff for maintenance, data quality, analytical and audit purposes. Paper records are stored securely in a locked office. Information available to the general public is of a statistical and non-identifiable nature. [more...]
References of Studies Using/Describing DB	10 most recent citations available in B.R.I.D.G.E.	10 most recent citations available in B.R.I.D.G.E.	10 most recent citations available in B.R.I.D.G.E.

CONCLUSIONS

Within 185 databases profiled, 22 were eligible for this breast cancer screening study. Although this analysis was done using databases currently in B.R.I.D.G.E., it indicates that more databases should consider including cost & laboratory data to facilitate CER studies.

This case study demonstrates how B.R.I.D.G.E. :
- Supports decision-making for database selection in CER studies;
- Serves as a useful tool to identify and compare health database attributes;
- Can be used as a teaching tool on healthcare databases; and
- Serves as a template to augment databases with more useful healthcare data.

This study was self-funded by DGI, LLC.

7th Asian Conference on Pharmacoepidemiology, Bengaluru, India. October 28, 2012

Comparing and Identifying Optimal Healthcare Databases for Comparative-Effectiveness of Breast Cancer Screening

A.J. Kapasi, S.A. Kamani, B. Nguyen-Khoa, V.P. Singh,
J. Tao, E.L. Goehring Jr, J.K. Jones

DGI, LLC, Arlington, VA, USA

Presented at the 7th Asian Conference on Pharmacoepidemiology. Bengaluru, India. October 28, 2012

DGI Center for Health Research and Education, LLC
1616 N Fort Myer Drive, Suite 1430 • Arlington, VA 22209 USA
www.bridgetodata.org T: +1 703 894 0804 info@bridgetodata.org