

Indexes, Registers, and Health Data Collection

Speaker: Heidi Jones

Indexes and registers (or registries) allow health information to be maintained and retrieved by health care facilities for the purpose of education, planning, and research.

The American Heritage Dictionary of the English Language provides the following definition for an index *“an index serves to guide, point out, or otherwise facilitate reference, especially an alphabetized list of names, places, and subjects treated in a print work, giving the page or pages on which the items is mentioned.”*

The American Heritage Dictionary of the English Language provides the following definition for a register *“a formal or official recording of items, names or actions.”*

Whereas, a registry is an organized system for the collection, storage, retrieval, analysis, and dissemination of information on individuals who have either a particular disease, a condition that predisposes to the occurrence of a health related event, or prior exposure to substance or circumstances known or suspect to cause adverse health effects.

As we previously noted in a prior unit, indexes, registers, and registries are considered secondary sources of patient information.

We will now discuss a very important index that is used in the health care setting. This index is referred to as the **Master Patient Index**. Rather than saying *Master Patient Index*, most health care professionals will refer to it as the “MPI.” The Master Patient Index is also referred to as the Master Person Index. The MPI is the key to locating patient records.

The MPI is retained permanently and filed alphabetically by the patient’s last name.

The demographic data that should be included on each patient’s MPI include the following:

- The patient’s name (their last name, first name, middle initial)
- The patient’s address (to include street, city, and zip code)
- Social Security number
- Their date of birth (make sure to use the mm/dd/yyyy format)
- The medical record number (this is assigned by the facility)
- The name of the facility and/or provider
- The type of care received (such as inpatient, outpatient, emergency, or physician’s office)

The main purpose of the MPI is to provide continuity of care. As previously mentioned, it is important to link the patient to their established medical record. The admission specialist needs to make sure to ask the patient if they have previously been seen in the facility. This process prevents the duplication of multiple medical record numbers for one patient. For the facility the MPI provides a customer database.

We will now discuss the disease, procedure, and physician indexes. Patient medical records are the source of data for the **indexes**. Data from the patient medical records abstracted into the computer system. The computer then generates a computerized database which creates the indexes.

The **disease index** is organized according to ICD-9-CM disease codes. The **procedure index** is organized according to ICD-9-CM and/or CPT/HCPCS procedures/services. The **physician index** is organized according to physician number. The physician number is assigned to the physician by the facility. Once a number has been assigned to a physician it is never reassigned to another physician.

The database for all three indexes also collects the following information:

- Patient demographic data (age, ethnicity, gender, inpatient admission and discharge dates or outpatient treatment date, and zip code)
- Financial data (third-party payer type and total charges)
- Medical data (attending physician, consulting physician, surgeon, medical service classification, disease and/or procedure/services code(s), date of surgery, and type of anesthesia)

If you are an HIT student, this is the information that you will abstract from the patient medical record to generate the indexes that we previously discussed.

Registers and registries contain information about a disease, such as cancer, or an event, such as a birth, and are maintained by individual health care facilities, federal, state government agencies, and private organizations.

Case report forms are submitted by health care facilities and provide reported data to sponsoring agencies, facilities, and organizations.

Common registers that are used in the hospital setting are the admission register, discharge register, and death register. Just as the name implies, these registers list the admissions to the hospital, the discharges from the hospital, and the deaths that occurred while hospitalized.

Registers and registries provide the facility, as well as the providers and public health officials, with information needed to assess and monitor the health given in a particular population.

There are many different types of registers and registries that are maintained in the United States. I am just going to mention a few as the list could go on and on:

- Birth Defects Registry
- Cancer Registry
- Implant Registry
- National Exposure Registry
- Organ Donor Registry
- United States Eye Injury Registry

Previously, I mentioned the task of **abstracting**. Registers receive data which have been abstracted from various organizations through the use of case abstracts.

Case abstracts have different data elements that must be collected. For example, the organ donor registry case abstract form will have different requirements than the implant registry. The data set used to collect and report data will depend on the requirements of the facility. Before the case abstract process can begin, a standard method of collecting and reporting individual data elements must be established so data can be easily compared. This is called a **data set**. In addition, the organization should develop a data set which is a collection of data element definitions.

The following is a list of the common data sets that are utilized in the collection of data for various facilities:

- Data Elements for Emergency Department Systems (DEEDS) is for use in hospital based emergency department
- Essential Medical Data Set (EMDS) also known as E-MEDS is also for emergency services
- Health Plan Employer Data and Information Set (HEDIS) is for managed care organizations
- Minimum Data Set (MDS) is for long term care facilities
- National Cancer Data Base (NCDB) is for hospital cancer registries
- Outcome and Assessment Information Set (OASIS) is for home health care agencies
- Uniformed Ambulatory Care Data Set (UACDS) is for ambulatory care facilities
- Uniform Clinical Data Set (UCDS) is for quality improvement organizations
- Uniform Hospital Discharge Data Set (UHDDS) is for hospitals. On a side note; the use of UHDDS is required for hospitals to participate in Medicare and Medicaid.
- Medical Information Bureau (MIB) is defined as a clearinghouse of medical information about people who apply for insurance
- National Practitioner Data Bank (NPDB) was legislated by the Health Care Quality and Improvement Act. This act was passed due to an increasing occurrence of medical malpractice and the need to improve the quality of medical care. The NPDB contains information about practitioners who engage in unprofessional behavior.

Vital statistics are compiled for births, deaths, fetal deaths, marriages, and divorces. The National Center for Health Statistics (NCHS) is a federal agency responsible for maintaining official vital statistics, while registration of vital events is a state function.

We'll now move onto descriptive Health Care Statistics –

Descriptive statistics summarize a set of data using charts, graphs, and tables. The primary purpose of statistics is to provide factual numerical information that can be used for decision making. There is a common phrase "a picture paints a thousand words." Graphic displays are commonly used in the hospital setting to display statistics. The HIM department is many times responsible for computing the facilities statistics as well as preparing the graphic displays.

Health care data is classified into the following categories:

- Aggregate data is based on performance and utilization and resource management
- Comparative data is used for health care services outcome measurement and research
- Patient-centric data directly relates to patient
- Transformed-based data is used for clinical and management decisions, support, and planning

The term **data quality** is a term that is used quite often in health care. Simply stated, inaccurate and incorrect data can result in serious and harmful errors. Data quality characteristics include the following:

- Data integrity; data has integrity if it is accurate, complete, consistent, up-to-date, and the same no matter where the data is recorded
- Data reliability; data is reliable if it is consistent throughout all systems in which it is stored, processed and/or retrieved
- Data validity; data is valid if it conforms to an expected range of values

The American Health Information Management Association (AHIMA) has defined four areas of Data Quality Management. They include the following:

- Data Application – purpose for which the data are collected
- Data Collection – processes by which data elements are accumulated
- Data Warehousing – processes and systems used to archive data and data journals
- Data Analysis – process of translating data into information utilized for an application

Continuous Quality Improvement (CQI) plays a role in data quality for the organization and should include data application, collection, warehousing, and analysis.

The characteristics of ensuring data quality includes the following:

- Data accessibility – ease with which data can be obtained
- Data accuracy – error free and correct
- Data comprehensiveness – presence of all required data elements in the patient record
- Data consistency – reliability of data regardless of the way in which data are stored, displayed, or processed
- Data definition – definite meanings and values of all elements so all present and future users understand the data
- Data granularity – definitions of each attribute and value of data at the correct level of detail
- Data precision – accurate data collection by defining expected data values
- Data relevancy – compilation of data that is valuable for the performance of a process or activity

Date timeliness or data currency – collection of up-to-date data and available to the user within a reasonable amount of time.