Health Care Data

Health Care Information Systems: A Practical Approach for Health Care Management
Karen A. Wager, Frances Wickham Lee, and John P. Glaser

Hye-Chung Kum (kum@tamu.edu)
Associate Professor
Population Informatics Lab (https://pinformatics.org/)
Course URL: http://pinformatics.org/phpm631

License: Health Information Technology by Hye-Chung Kum is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

Learning Objectives

- Health Care Data and Information Defined
  - What are health data and health information?
- Health Care Data and Information Sources
  - Where does health data originate and why?
  - When does health care data become health care information?
- Health Care Data Uses
  - How do health care organizations use data?
  - What is the impact of the trend toward analytics and big data on health care data?
- Health Care Data Quality
  - How does the quality of health data impact its use?

Definitions

- Health Care Data vs. Health Care Information
  - Data need to be processed to be information
  - On a continuum, not clear cut
  - Collected by a covered entity
  - One person’s data maybe be another person’s information depending on the purpose

- Electronic Medical Records (EMR) vs. Electronic Health Records (EHR)

Purposes: Patient Records

- Patient care
- Communication
- Legal documentation
- Billing and reimbursement
- Research and quality management
- Population health
- Public health

Components: Patient Records

- Identification screen
- Problem list
- Medication record
- History and physical
- Progress notes (SOAP)
  - Subjective findings, Objective Findings, Assessment, Plan
- Consultation: outside provider
- Physician’s orders
- Imaging and x-ray reports
- Laboratory reports
- Consent & authorization forms
- Operative reports
- Pathology reports
- Discharge summary

Components: Claims

- Accounting/Billing
  - Verify insurance coverage
  - Bill third party payers
    - UB-04 (CMS-1450 (B37): institutional (hospitals, SNFs, home health, FQHC, community mental health centers etc.)
    - CMS-1500 (837P): non-institutional provider claims (private physician services)
  - AMA
  - Process payments upon receipt
**EHR Information Screen**

![EHR Information Screen](Source Medical University of South Carolina EHR)

**EHR Problem List**

![EHR Problem List](Source Medical University of South Carolina EHR)

**EHR Progress Notes**

![EHR Progress Notes](Source Medical University of South Carolina EHR)

**EHR Lab Report**

![EHR Lab Report](Source Medical University of South Carolina EHR)

**Codes: Diagnostic and Procedural**

- ICD-10 (International Classification of Diseases)
  - ICD-10-CM (clinical modification) vs ICD-10-PCS (procedure coding system)
- CPT (Current Procedural Terminology)
  - Copyrighted with all rights to publication and distribution held by the AMA
  - Provides a uniform language for describing medical and surgical services
  - HCPCS level 1: standard for physician's office, outpatient, ambulatory care
- Coding plays a major role in reimbursement for care
  - Both up coding and down coding is a problem

**Accountable Care Act**

- Shifting focus from episodic care to population health
- Successful population health require extensive coordination of care across providers and community organizations
  - Comprehensive shared care plan (CSCP)
  - Care managers are needed to interact with patients on a regular basis during and in between clinical encounters
- Reliance on HIT
Data Analysis

Four basic elements of data analysis
- Source of data
  - EHR, claims data, laboratory data, etc.
- Stored in a retrievable manner
  - Database or data warehouse
- Analytical tool applied
  - Mathematical statistics, probability models, predictive models, etc.
- Reported in a usable manner

Outline: Health Care Data Use

- Operational vs Decision Support Systems
- What is Data Science/Business Intelligence
  - What is Data Science?
  - What is Big Data?
  - Overview of Data Mining
- Understanding Data

Operational vs Decision Support Systems

- Operational Systems
  - Support day to day transactions
  - Contain current, "up to date" data
  - Examples: EMR, customer orders, inventory levels, bank account balances
- Decision Support Systems
  - Support strategic decision making
  - Contain historical, “summarized” data
  - Examples:
    - Clinical support: what treatment is best?
    - Population health:
    - Management support: performance summary, market segmentation

Operational Application: EMR

DSS (Reports)

- PricePoint
  - for consumers
  - http://www.txpricepoint.org/

Video

- TX Mental Health Landscape (2:46)
  - https://www.youtube.com/watch?v=8dPqQt0yXJA
- Wealth Inequality (1:30)
  - https://www.youtube.com/watch?v=QPKKQnimsM
- Survey
Outline: Health Care Data Use

* Operational vs. Decision Support Systems
* What is Data Science/Business Intelligence
  * What is Data Science?
  * What is Big Data?
  * Overview of Data Mining
* Understanding Data

What is Data Science?

* Other words
  * Knowledge Discovery & Data mining (KDD)
  * Business Intelligence / Business Analytics
  * Collecting and refining information from many sources
  * Analyzing and presenting the information in useful ways
  * So people can make better business decisions

Data Science

Knowledge Discovery & Data mining (KDD)

Big Data: operational data  KDD: Clean, Merge, Reprocess

Human consumable, valid, novel, potentially useful, and ultimately understandable information

The Virtuous Cycle of Data to Decision & Action

1. Identify the business problem
2. Transform data into information using data mining techniques
3. Act on the information
4. Measure the results

Data to Decision

Data Science Definition (Big Data - less consensus)

Data Science is the extraction of actionable knowledge directly from data through a process of discovery, hypothesis, and analytical hypothesis analysis.

A Data Scientist is a practitioner who has sufficient knowledge of the overlapping regimes of expertise in business needs, domain knowledge, analytical skills and programming expertise to manage the end-to-end scientific method process through each stage in the big data lifecycle.
KDD Process

- Data cleaning & integration
- Feature Selection (what vars?)
- Analysis / Data mining
- Validation / Evaluation
- Action

Thomas Davenport: Competing on Analytics

- Skill set for good data scientists
  - IT & Programming skills
  - Statistical skills
  - Business skills:
    - Understand pros/cons of decisions & actions
    - Communication skills
    - Excel / PowerPoint
  - Intense curiosity: the most important skill or trait. “A desire to go beyond the surface of a problem, find the question at its heart, and distill them into a very clear set of hypotheses that can be tested.”

Data science teams need people with the skills and curiosity to ask the big questions (oreilly)

- Technical expertise: the best data scientists typically have deep expertise in some scientific discipline.
- Curiosity: a desire to go beneath the surface and discover and distill a problem down into a very clear set of hypotheses that can be tested.
- Storytelling: the ability to use data to tell a story and to be able to communicate it effectively.
- Cleverness: the ability to look at a problem in different, creative ways.
- Health is a very important domain
  - Team lead: good questions, good interpretation & implications

Job market of data scientists

- Statisticians will be the next sexy job
  - Google Chief Economist Hal Varian
  - Shortage of 190,000 data scientists by the year 2019
    - McKinsey Global Institute

New Era in Science: Big Data Science

- Data is the new raw material of business: an economic input almost on par with capital and labor. (Microsoft’s Craig Mundie)
- Those who can harness the power of data will lead the next century and drive innovation in commerce, scientific discovery, healthcare, finance, energy, government, and countless other fields.
- Students who learn to be a data science will be in high demand.

Outline: Health Care Data Use

- Operational vs. Decision Support Systems
- What is Data Science/Business Intelligence
  - What is Data Science?
  - What is Big Data?
  - Overview of Data Mining
- Understanding Data
Properties of BIG DATA: 4V

- Volume: constantly generating
- Velocity: constantly changing
- Variety: expressed in many ways
- Veracity: lots of errors

(Value)

EXAMPLE: the INTERNET!
What do you do to find information/knowledge on the Internet?

The Big Data Problem: Nutshelled
Michael Franklin (UC Berkley)

Something's gotta give:

Money

Quality (precision)

Time

AMP Lab: Integrating Three Key Resources

- Machine Learning, Statistical Methods
- Prediction, Business Intelligence
- Clusters and Clouds
- Warehouse Scale Computing
- Crowdsourcing, Human Computation
- Data Scientists, Analysts

Outline: Health Care Data Use

- Operational vs. Decision Support Systems
- What is Data Science/Business Intelligence
  - What is Data Science?
  - What is Big Data?
  - Overview of Data Mining
- Understanding Data

What is Data Mining?

- Using a combination of artificial intelligence, machine learning, and statistical analysis to analyze data
- and discover useful patterns that are "hidden" there

Sample Applications

- Direct Marketing
  - Identify which prospects should be included in a mailing list
  - Identify that encourage direct purchase
- Market segmentation
  - Identify common characteristics of customers who buy same products
  - Profile common characteristics in homogeneous patient group


- Customer churn
  - Predict which customers are likely to leave your company for a competitor
- Potentially Preventable Readmissions

- Market basket analysis
  - Identify what products are likely to be bought together
  - Care coordination: common services for a condition (bundled services)

- Insurance claims analysis
  - Discover patterns of fraudulent transactions (medical fraud)
  - Compare current transactions against these patterns
Business uses of data mining: Essentially five tasks

- **Classification**
  - Classify credit applicants as low, medium, high risk
  - Classify insurance claims as normal, suspicious
- **Estimation**
  - Estimate the probability of a direct mailing response
  - Estimate the potential cohort size for a clinical trial
- **Prediction**
  - Predict which customers will leave within six months
  - Predict which patient will return to the ED
- **Affinity Grouping**
  - Find out what books to recommend to Amazon.com users
  - Find treatment regime that was successful for similar patient
- **Description**
  - Help understand large volumes of data by uncovering interesting, useful, and actionable patterns

Applications in Health

- A March 2014 poll from MeriTalk and EMC found that 63 percent of healthcare executives in the federal government believe that big data will improve population health management.
- **Examples**
  - Manage population health
  - Accountable Care Organizations (ACO)
  - Clinical decision support
  - Cohort identification for clinical trials
  - Medical fraud detection

Outline: Health Care Data Use

- Operational vs. Decision Support Systems
- **What is Data Science/Business Intelligence**
  - What is Data Science?
  - What is Big Data?
  - Overview of Data Mining
- **Understanding Data**

Healthcare Data Quality

- Depends on the use of the data
- Traditionally (file cabinet):
  - Patient Clinical/Claim Records
    - Episode
    - Generally from a single organization
- Today (conveyor belt - big data):
  - EHR/Electronic Claims Record
    - Continuous instead of episodic
    - Criteria for quality has shifted

AHIMA

- Accuracy
- Accessibility
- Comprehensive
  - all relevant without inundating
- Consistency
- Currency
  - Data go obsolete
- **Clear Definition**
  - Used and understood to be the same
- **Granularity:** unit of data
- **Precision**
- **Relevancy**
- **Timeliness**

Weiskopf and Weng

- Completeness
- Correctness: free of error
- Concordance: consistency, reliability
- Plausibility: believability, validity
- Currency
Summary

- Value of Health Care Data and Information
- Importance of Health Care Data and Information Sources
- Uses and Analysis of Health Care Data
- Assurance of Health Care Data Quality