Big Data for Rehabilitation: Promises, Pitfalls, and Future Potential

Kenneth J. Ottenbacher
University of Texas Medical Branch
1. Introduction
2. Challenges & Opportunities
3. Resources – Center for Large Data Research & Data Sharing (CLDR)
4. Questions
The march of quantification, made possible by enormous new sources of data, will sweep through academia, business, health care and government. There is no area that is going to be untouched.”

Science 2011
The Challenges

Large datasets can be overwhelming
Rehabilitation investigators are trained to conduct (prospective) clinical, patient-oriented research...

not comprehensive or secondary analysis of large data.
The Opportunities

- Advances in bioinformatics, statistical computing, information technology and the internet have resulted in access to large amounts of data.
- Health and biomedical data are available that include information related to rehabilitation, recovery, and disability.
NIH Big Data to Knowledge (BD2K)

Areas of BD2K

1. Facilitating Broad Use of Data
2. Analysis Methods and Software
3. Enhancing Training
4. Centers of Excellence

https://datascience.nih.gov/bd2k/about
What is “Big Data”?  

NIH BD2K Definition

“Big Data refers to the complexity, challenges, and new opportunities presented by the combined analysis of data. In biomedical research, these data sources include the diverse, complex, disorganized, massive, and multimodal data being generated by researchers, hospitals, and mobile devices around the world.”

https://datascience.nih.gov/bd2k/about/what
Characteristics of “Big Data”?

Includes everything from Twitter to the Genome
Large Data vs. Big Data

Adapted from: Roski J et al. Health Aff 2014;33:1115-1122

Big Data

- Analyzes different data types on massive scale resulting in predictive and real-time analyses
- Large (or all) amounts of data (gigabytes to petabytes)
- Accurate historical observations
- Predictive and real-time analytics capabilities

Large Data Analytics

- Includes advanced statistical approaches involving sophisticated software
- Data of same form
- Mega-bytes to tera-bytes

Basic Analytics

- Includes traditional observational and software-based statistical approaches
- Data of same form
- Mega-bytes to tera-bytes

Big Data Computing

- Data consolidated, allows analytical work to be streamlined
Personal Motion Technology

Continuous recording of biological or performance information generates large amounts of data.

Opportunities (Example)
Opportunities (Example)

PCORnet, the National Patient-Centered Clinical Research Network, is an initiative of PCORI. It is designed to make it faster, easier, and less costly to conduct clinical research than is now possible by using the power of large health data and patient partnerships.

http://www.pcornet.org/
What we have done … Link research to Federal healthcare legislation & priorities:

1. ACA – Hospital Readmission Reduction Program

2. ACA – Bundled Payment for Care Improvement
   • Comprehensive Care for Joint Replacement

3. ACA – Value-based Healthcare - Quality Measure Reporting

4. MedPAC Reports to Congress
Hospital Readmission is a National Quality Indicator for Healthcare Reform

We examined national data from 1.7 million Medicare files to determine rates of hospital readmission for the six largest impairment groups receiving inpatient medical rehabilitation in the U.S.
Activities - CLDR

- Web Site  
  [http://rehabsciences.utmb.edu/cldr](http://rehabsciences.utmb.edu/cldr)

- Education & Training

- Data Directory

- Pilot Projects

- Visiting Scholars Program

- Data Archiving & Sharing Program
Education & Training

CLDR education and training activities cover three fundamental aspects of using large data for rehabilitation research: 1) Identifying and accessing large data sources; 2) Managing and analyzing large datasets; and, 3) Archiving and sharing research data with other investigators.

CoHSTAR Summer Institute on Health Services Research at Boston University - Videos Available
On June 29, 2016 CoHSTAR held its first Summer Institute on Health Services Research at Boston University. Summer Institute sessions were webinar and the archived proceedings are now available for viewing, including a presentation by CLDR Director Ken Ottenbacher, PhD, OTR. Access to the archived webinar proceedings is free with registration.

CLDR Institute on Large Data and Data Sharing Research: Issues, Methods & Opportunities - Videos Available
CLDR Director Ken Ottenbacher, PhD, OTR, and colleagues offered a pre-conference Institute on Large Data and Data Sharing Research: Issues, Methods & Opportunities co-sponsored by the American Occupational Therapy Foundation (AOTF) and American Occupational Therapy Association (AOTA) at AOTA’s national conference in Chicago on April 6th, 2016.
Pilot Projects

Funding Opportunities

The CLDR’s pilot studies program is designed to offer practical experience in managing, analyzing, and interpreting findings using large data analytics and secondary data analysis. The Center will fund pilot studies in two categories: Category 1 will include research examining questions using a large administrative or healthcare research dataset and Category 2 will involve the archiving of existing data from a completed study.

Application Process & Eligibility Criteria

- Category 1: Large Data Pilot Studies
- Category 2: Data Sharing and Archiving Pilot Studies
- Current & Completed Projects

Important Dates for Category 1 Projects

- March 1: Letter of Intent deadline
- March 15: Notifications in response to LOI
- April 15: Pilot application deadline
- May 15: Funding decision notifications
- July 1: Start Date

Note: Category 2 project applications are accepted year round.

Submit a Pilot Application
Category 1: Large Data Pilot Studies

These include traditional pilot studies examining a research question using large data (e.g., survey or administrative dataset) that address an issue related to rehabilitation, recovery, or disability. Funding up to $25,000
Category 2: Data Sharing and Archiving Pilot Studies

Category 2 projects are focused on archiving data from a completed rehabilitation study with support being given to either the PI or a member of their research team.

- Funding up to $10,000
- Applications accepted year-round
Visiting Scholars Program

Investigators can be supported for periods of up to six months at one of the consortium institutions to work with experienced investigators using large datasets.
New CLDR activity related to federal mandates for data sharing

White House Office of Science and Technology Policy (February 22, 2013)

DASH is a resource to store and access de-identified data from NICHD-funded research studies for secondary research use.

https://dash.nichd.nih.gov/
FITBIR functions as a repository for data from TBI studies and supports a free web-based application with an interface that allows investigators to contribute and access data.

https://fitbir.nih.gov/
Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?

T.S. Eliot (The Rock)

Where is the information we have lost in data?
Rehabilitation Sciences
Academic Division &
Research Center

Phone: (409) 747-1637
Email: rehab.info@utmb.edu

http://rehabsciences.utmb.edu/cldr

Information
Disclosures (PI or core leader)                                                                                       Anne Deutsch, RN, PhD, CRRN
• **Research**                                                                                                       Brian Downer, PhD
  P2C-HD065702 (PI, NIH)                                                                                               Karl Eschbach, PhD
  R01-HD069443 (PI, NIH)                                                                                              James Goodwin, MD
  R01 MD-010355 (PI, NIH)                                                                                             James Graham, PhD
  H133G140127 (PI, NIDILRR)                                                                                            Ickpyo Hong, PhD, OTR*
  R24-HS022134 (Core, AHRQ)                                                                                            Bret Howrey, PhD
  P30-AG024832 (Core, NIH)                                                                                             Jessica Jarvis, MS*
• **Training**                                                                                                         Amol Karmarkar, PhD, MPH
  K12-HD055929 (PI, NIH)                                                                                               Amit Kumar, PT, PhD, MPH*
  H133P110012 (PI, NIDILRR)                                                                                            Yong-Fang Kuo, PhD
  Chih-ying (Cynthia) Li, PhD, OTR*                                                                                   Yi-Li Lin, MS
  *Postdoc or Graduate Student                                                                                         Addie Middleton, PT, DPT, PhD
  Tim Reistetter, PhD, OTR                                                                                              Jie Zhou, MS

---

**Large Data Research Team**

Soham Al Snih, MD, PhD
Winston Chan, PhD
Nai-Wei Chen, PhD

*Postdoc or Graduate Student*