Caveats and Recommendations for Use of Operational Electronic Health Record Data for Research and Quality Measurement

William Hersh, MD, FACP, FACMI
Diplomate, Clinical Informatics, American Board of Preventive Medicine
Professor and Chair
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
Portland, OR, USA
Email: hersh@ohsu.edu
Web: www.billhersh.info
Blog: http://informaticsprofessor.blogspot.com

References

Altman, RB (2012). Translational bioinformatics: linking the molecular world to the clinical world. *Clinical Pharmacology and Therapeutics*. 91: 994-1000.
Amarasingham, R, Patel, PC, et al. (2013). Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. *BMJ Quality & Safety*. 22: 998-1005.
Barkhuysen, P, deGrauw, W, et al. (2014). Is the quality of data in an electronic medical record sufficient for assessing the quality of primary care? *Journal of the American Medical Informatics Association*. 21: 692-698.
Bayley, KB, Belnap, T, et al. (2013). Challenges in using electronic health record data for CER: experience of four learning organizations. *Medical Care*. 51: S80-S86.
Blumenthal, D (2011). Implementation of the federal health information technology initiative. *New England Journal of Medicine*. 365: 2426-2431.
Blumenthal, D (2011). Wiring the health system--origins and provisions of a new federal program. *New England Journal of Medicine*. 365: 2323-2329.
Butler, D (2013). When Google got flu wrong. *Nature*. 494: 155-156.
Chapman, WW, Christensen, LM, et al. (2004). Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artificial Intelligence in Medicine*. 33: 31-40.
Charles, D, Gabriel, M, et al. (2014). Adoption of Electronic Health Record Systems among U.S. Non-federal Acute Care Hospitals: 2008-2013. Washington, DC, Department of Health and Human Services. http://www.healthit.gov/sites/default/files/oncdatabrief16.pdf
Collins, FS, Hudson, KL, et al. (2014). PCORnet: turning a dream into reality. *Journal of the American Medical Informatics Association*: Epub ahead of print.
Crosslin, DR, McDavid, A, et al. (2012). Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. *Human Genetics*. 131: 639-652.
Dahabreh, IJ and Kent, DM (2014). Can the learning health care system be educated with observational data? *Journal of the American Medical Association*. 312: 129-130.
Danaei, G, Rodríguez, LA, et al. (2011). Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical Methods in Medical Research*. 22: 70-96.

deLusignan, S and vanWeel, C (2005). The use of routinely collected computer data for research in primary care: opportunities and challenges. *Family Practice*. 23: 253-263.
Demner-Fushman, D, Abhyankar, S, et al. (2012). NLM at TREC 2012 Medical Records Track. *The Twenty-First Text REtrieval Conference Proceedings (TREC 2012)*, Gaithersburg, MD. National Institute for Standards and Technology http://trec.nist.gov/pubs/trec21/papers/NLM.medical.final.pdf
Denny, JC, Crawford, DC, et al. (2011). Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical records for genome- and phenome-wide studies. *American Journal of Human Genetics*. 89: 529-542.
Denny, JC, Ritchie, MD, et al. (2010). Identification of genomic predictors of atrioventricular conduction: using electronic medical records as a tool for genome science. *Circulation*. 122: 2016-2021.
Detmer, DE and Shortliffe, EH (2014). Clinical informatics: prospects for a new medical subspecialty. *Journal of the American Medical Association*. 311: 2067-2068.
Diamond, CC, Mostashari, F, et al. (2009). Collecting and sharing data for population health: a new paradigm. *Health Affairs*. 28: 454-466.
Edinger, T, Cohen, AM, et al. (2012). Barriers to retrieving patient information from electronic health record data: failure analysis from the TREC Medical Records Track. *AMIA 2012 Annual Symposium*, Chicago, IL. 180-188.
Fleurence, RL, Curtis, LH, et al. (2014). Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association*: Epub ahead of print.
Gerbier, S, Yarovaya, O, et al. (2011). Evaluation of natural language processing from emergency department computerized medical records for intra-hospital syndromic surveillance. *BMC Medical Informatics & Decision Making*. 11: 50. http://www.biomedcentral.com/1472-6947/11/50
Gildersleeve, R and Cooper, P (2013). Development of an automated, real time surveillance tool for predicting readmissions at a community hospital. *Applied Clinical Informatics*. 4: 153-169.
Ginsberg, J, Mohebbi, MH, et al. (2009). Detecting influenza epidemics using search engine query data. *Nature*. 457: 1012-1014.
Gottesman, O, Kuivaniemi, H, et al. (2013). The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genetics in Medicine*. 15: 761-771.
Greene, SM, Reid, RJ, et al. (2012). Implementing the learning health system: from concept to action. *Annals of Internal Medicine*. 157: 207-210.
Hebert, C, Shivade, C, et al. (2014). Diagnosis-specific readmission risk prediction using electronic health data: a retrospective cohort study. *BMC Medical Informatics & Decision Making*. 14: 65. http://www.biomedcentral.com/1472-6947/14/65
Henning, KJ (2004). What is syndromic surveillance? *Morbidity and Mortality Weekly Report*. 53(Suppl): 5-11. http://www.cdc.gov/mmwr/preview/mmwrhtml/su5301a3.htm
Hersh, W (2010). The health information technology workforce: estimations of demands and a framework for requirements. *Applied Clinical Informatics*. 1: 197-212.
Hersh, WR, Cimino, JJ, et al. (2013). Recommendations for the use of operational electronic health record data in comparative effectiveness research. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 1: 14. http://repository.academyhealth.org/egems/vol1/iss1/14/

Hersh, WR, Weiner, MG, et al. (2013). Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical Care*. 51(Suppl 3): S30-S37.
Horner, P and Basu, A (2012). Analytics & the future of healthcare. Analytics, January/February 2012. http://www.analytics-magazine.org/januaryfebruary-2012/503-analytics-a-the-future-of-healthcare
Hsiao, CJ and Hing, E (2014). Use and Characteristics of Electronic Health Record Systems Among Office-based Physician Practices: United States, 2001–2013. Hyattsville, MD, National Center for Health Statistics. http://www.cdc.gov/nchs/data/databriefs/db143.htm
Jha, AK, Joynt, KE, et al. (2012). The long-term effect of Premier pay for performance on patient outcomes. *New England Journal of Medicine*: Epub ahead of print.
King, B, Wang, L, et al. (2011). Cengage Learning at TREC 2011 Medical Track. *The Twentieth Text REtrieval Conference Proceedings (TREC 2011)*, Gaithersburg, MD. National Institute for Standards and Technology
Kullo, LJ, Ding, K, et al. (2010). A genome-wide association study of red blood cell traits using the electronic medical record. *PLoS ONE*. 5(9): e13011.
Lazer, D, Kennedy, R, et al. (2014). Big data. The parable of Google Flu: traps in big data analysis. *Science*. 343: 1203-1205.
Lee, SJ and Walter, LC (2011). Quality indicators for older adults: preventing unintended harms. *Journal of the American Medical Association*. 306: 1481-1482.
MacKenzie, SL, Wyatt, MC, et al. (2012). Practices and perspectives on building integrated data repositories: results from a 2010 CTSA survey. *Journal of the American Medical Informatics Association*. 19(e1): e119-e124.
Miller, DR, Safford, MM, et al. (2004). Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. *Diabetes Care*. 27(Suppl 2): B10-21.
Mirnezami, R, Nicholson, J, et al. (2012). Preparing for precision medicine. *New England Journal of Medicine*. 366: 489-491.
Newton, KM, Peissig, PL, et al. (2013). Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. *Journal of the American Medical Informatics Association*. 20(e1): e147-154.
Parsons, A, McCullough, C, et al. (2012). Validity of electronic health record-derived quality measurement for performance monitoring. *Journal of the American Medical Informatics Association*. 19: 604-609.
Pathak, J, Bailey, KR, et al. (2013). Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPn consortium. *Journal of the American Medical Informatics Association*. 20: e341-e348.
Rea, S, Pathak, J, et al. (2012). Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: The SHARPn project. *Journal of Biomedical Informatics*. 45: 763-771.
Richesson, RL, Rusincovitch, SA, et al. (2013). A comparison of phenotype definitions for diabetes mellitus. *Journal of the American Medical Informatics Association*. 20: e319-e326.
Safran, C, Bloomrosen, M, et al. (2007). Toward a national framework for the secondary use of health data: an American Medical Informatics Association white paper. *Journal of the American Medical Informatics Association*. 14: 1-9.

Serumaga, B, Ross-Degnan, D, et al. (2011). Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *British Medical Journal*. 342: d108. http://www.bmj.com/content/342/bmj.d108
Tannen, RL, Weiner, MG, et al. (2008). Replicated studies of two randomized trials of angiotensin-converting enzyme inhibitors: further empiric validation of the 'prior event rate ratio' to adjust for unmeasured confounding by indication. *Pharmacoepidemiology and Drug Safety*. 17: 671-685.
Tannen, RL, Weiner, MG, et al. (2009). Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings. *British Medical Journal*. 338: b81. http://www.bmj.com/cgi/content/full/338/jan27_1/b81
Tannen, RL, Weiner, MG, et al. (2007). A simulation using data from a primary care practice database closely replicated the Women's Health Initiative trial. *Journal of Clinical Epidemiology*. 60: 686-695.
Voorhees, E and Hersh, W (2012). Overview of the TREC 2012 Medical Records Track. *The Twenty-First Text REtrieval Conference Proceedings (TREC 2012)*, Gaithersburg, MD. National Institute of Standards and Technology http://trec.nist.gov/pubs/trec21/papers/MED12OVERVIEW.pdf
Wang, TY, Dai, D, et al. (2011). The importance of consistent, high-quality acute myocardial infarction and heart failure care results from the American Heart Association's Get with the Guidelines Program. *Journal of the American College of Cardiology*. 58: 637-644.
Wei, WQ, Leibson, CL, et al. (2013). The absence of longitudinal data limits the accuracy of high-throughput clinical phenotyping for identifying type 2 diabetes mellitus subjects. *International Journal of Medical Informatics*. 82: 239-247.
Weiner, MG, Barnhart, K, et al. (2008). Hormone therapy and coronary heart disease in young women. *Menopause*. 15: 86-93.
Xu, H, Aldrich, MC, et al. (2014). Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. *Journal of the American Medical Informatics Association*: Epub ahead of print.
Zhang, Z and Sun, J (2010). Interval censoring. *Statistical Methods in Medical Research*. 19: 53-70.

Caveats and Recommendations for Use of Operational Electronic Health Record Data for Research and Quality Measurement

William Hersh, MD, FACP, FACMI
Diplomate, Clinical Informatics, ABPM
Professor and Chair
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
Portland, OR, USA
Email: hersh@ohsu.edu
Web: www.billhersh.info
Blog: http://informaticsprofessor.blogspot.com

Disclosures

- William Hersh, MD
  – Has no financial relationships to disclose
  – Will not be discussing off-label/investigative use(s) of commercial devices
- Planning committee members have nothing to disclose: Vitaly Herasevich, MD, PhD; Christopher Chute, MD; Brian Pickering, MD; and Ms. Robin Williams

Outline

- Opportunities for secondary use or re-use of clinical data for research and other purposes
- Caveats of using operational clinical data
- Recommendations for using operational clinical data
- Research project: information retrieval of medical records for cohort discovery

OREGON HEALTH & SCIENCE UNIVERSITY

US has made substantial investment in health information technology (HIT)

"To improve the quality of our health care while lowering its cost, we will make the immediate investments necessary to ensure that within five years, all of America's medical records are computerized … It just won't save billions of dollars and thousands of jobs – it will save lives by reducing the deadly but preventable medical errors that pervade our health care system."

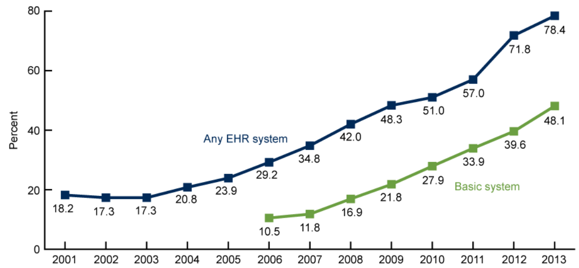
January 5, 2009

Health Information Technology for Economic and Clinical Health (HITECH) Act of the American Recovery and Reinvestment Act (ARRA) (Blumenthal, 2011)

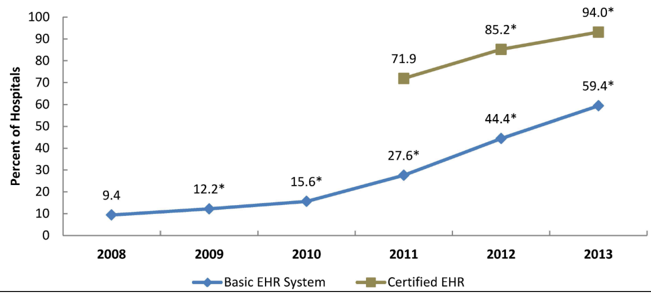
- Incentives for electronic health record (EHR) adoption by physicians and hospitals (up to $27B)
- Direct grants administered by federal agencies ($2B, including $118M for workforce development)

OREGON HEALTH & SCIENCE UNIVERSITY

Which has led to significant EHR adoption in the US



(Hsiao, 2014)



(Charles, 2014)

Providing opportunities for "secondary use" or "re-use" of clinical data

- (Safran, 2007; SHARPn, Rea, 2012)
- Using data to improve care delivery – predictive analytics
- Healthcare quality measurement and improvement
- Clinical and translational research – generating hypotheses and facilitating research
- Public health surveillance – including for emerging threats
- Implementing the learning health system

6

OREGON HEALTH & SCIENCE UNIVERSITY

Using data to improve healthcare

- With shift of payment from "volume to value," healthcare organizations will need to manage information better to provide better care (Diamond, 2009; Horner, 2012)
- Predictive analytics is use of data to anticipate poor outcomes or increased resource use –applied by many to problem of early hospital re-admission (e.g., Gildersleeve, 2013; Amarasingham, 2013; Herbert, 2014)
- A requirement for "precision medicine" (Mirnezami, 2012) and "personalized medicine" (Altman, 2012)

OREGON HEALTH & SCIENCE UNIVERSITY

Quality measurement and improvement

- Quality measures increasingly used in US and elsewhere to make care more "accountable"
  - Used more for process than outcome measures (Lee, 2011), e.g., Stage 1 meaningful use
- In UK, pay for performance schemes achieved early value but fewer further gains (Serumaga, 2011)
- In US, some quality measures found to lead to improved patient outcomes (e.g., Wang, 2011), others not (e.g., Jha, 2012)
- Desire is to derive automatically from EHR data, but this has proven challenging with current systems (Parsons, 2012; Pathak, 2013; Barkhuysen, 2014)

OREGON HEALTH & SCIENCE UNIVERSITY

Clinical and translational research

- Led in part by activities of NIH Clinical and Translational Science Award (CTSA) Program (Mackenzie, 2012)
- One of largest and most productive efforts has been eMERGE Network – connecting genotype-phenotype (Gottesman, 2013; Newton, 2013)
  - http://emerge.mc.vanderbilt.edu
  - Has used EHR data to identify genomic variants associated with atrioventricular conduction abnormalities (Denny, 2010), red blood cell traits (Kullo, 2010), white blood cell count abnormalities (Crosslin, 2012), thyroid disorders (Denny, 2011), etc.

9

Clinical and translational research (cont.)

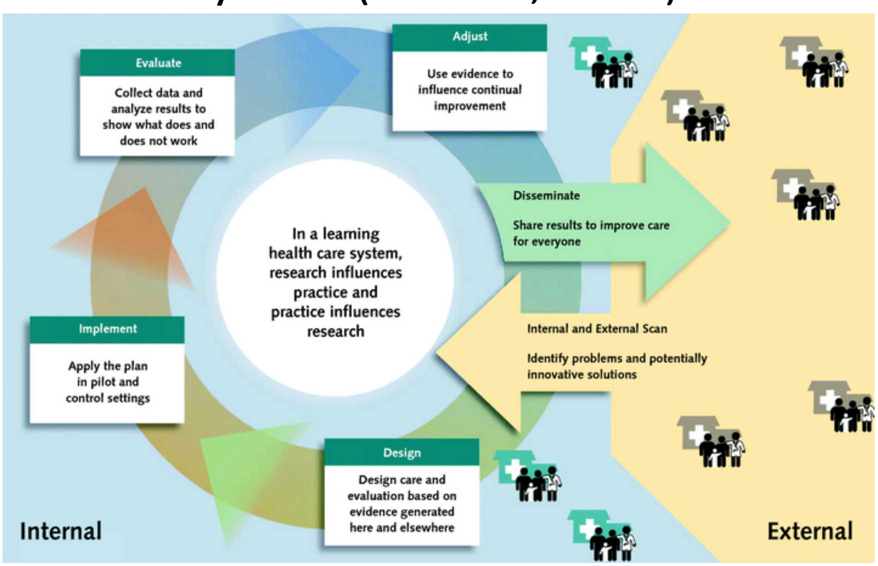
- Other successes include replication of clinical studies, e.g.,
  - Randomized controlled trials (RCT)
    - Women's Health Initiative (Tannen, 2007; Weiner, 2008)
    - Other cardiovascular diseases (Tannen, 2008; Tannen, 2009) and value of statin drugs in primary prevention of coronary heart disease (Danaei, 2011)
  - Observational studies
    - Metformin and reduced cancer mortality rate (Xu, 2014)
- Much potential for using propensity scores with observational studies as complement to RCTs
  - Often but not always obtain same results as RCTs (Dahabreh, 2014)

10

Public health

- "Syndromic surveillance" aims to use data sources for early detection of public health threats, from bioterrorism to emergent diseases
- Interest increased after 9/11 attacks (Henning, 2004; Chapman, 2004; Gerbier, 2011)
- Ongoing effort in Google Flu Trends
  - http://www.google.org/flutrends/
  - Search terms entered into Google predicted flu activity but not early enough to intervene (Ginsberg, 2009)
  - Performance in recent years has been poorer (Butler, 2013)
  - Case of needing to avoid "Big Data hubris" (Lazer, 2014)

Implementing the learning healthcare system (Greene, 2012)



Caveats for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research

William R. Hersh, MD, Mark G. Weiner, MD,† Peter J. Embi, MD, MS,‡ Judith R. Logan, MD, MS,* Philip R.O. Payne, PhD,‡ Elmer V. Bernstam, MD, MSE,§ Harold P. Lehmann, MD, PhD,‖ George Hripcsak, MD, MS,¶ Timothy H. Hartzog, MD, MS,# James J. Cimino, MD,** and Joel H. Saltz, MD, PhD††*

Operational clinical data may be (Medical Care, 2013):

- Inaccurate
- Incomplete
- Transformed in ways that undermine meaning
- Unrecoverable for research
- Of unknown provenance
- Of insufficient granularity
- Incompatible with research protocols

Inaccurate

- Documentation not always a top priority for busy clinicians (de Lusignan, 2005)
- Analysis of EHR systems of four known national leaders assessed use of data for studies on treatment of hypertension and found five categories of reasons why data were problematic (Bayley, 2013)
  - Missing
  - Erroneous
  - Un-interpretable
  - Inconsistent
  - Inaccessible in text notes

OREGON HEALTH & SCIENCE UNIVERSITY

Incomplete

- Not every diagnosis is recorded at every visit; absence of evidence is not always evidence of absence, an example of a concern known by statisticians as *censoring* (Zhang, 2010)
- Makes tasks such as identifying diabetic patients challenging (Miller, 2004; Wei, 2013; Richesson, 2013)
- Undermine ability to automate quality measurement
  - Measures under-reported based on under-capture of data due to variation in clinical workflow and documentation practices (Parsons, 2012)
  - Correct when present but not infrequently missing in primary care EHRs (Barkhuysen, 2014)

OREGON HEALTH & SCIENCE UNIVERSITY

"Idiosyncrasies" of clinical data (Hersh, 2013)

- "Left censoring" – First instance of disease in record may not be when first manifested
- "Right censoring" – Data source may not cover long enough time interval
- Data might not be captured from other clinical (other hospitals or health systems) or non-clinical (OTC drugs) settings
- Bias in testing or treatment
- Institutional or personal variation in practice or documentation styles
- Inconsistent use of coding or standards

OREGON HEALTH & SCIENCE UNIVERSITY

Overcoming the caveats: recommendations for EHR data use

- (Hersh, 2013)
- Assessing and using data
- Adaptation of “best evidence” approaches to use of operational data
- Need for standards and interoperability
- Appropriate use of informatics expertise


AcademyHealth

eGEMs

Home > EDM_PUBLICATIONS > eGEMs > Vol. 1 (2013) > Iss. 1

Recommendations for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research

Download

William R. Hersh, Oregon Health & Science University — Follow
James Cimino, National Institutes of Health Clinical Center — Follow
Philip R.O. Payne, The Ohio State University — Follow
Peter Embi, The Ohio State University — Follow
Judith Logan, Oregon Health & Science University — Follow
Mark Weiner, AstraZeneca — Follow
Elmer V. Bernstam, The University of Texas Health Science Center at Houston — Follow
Harold Lehmann, Johns Hopkins University — Follow
George Hripcsak, Columbia University — Follow
Timothy Hartzog, Medical University of South Carolina — Follow
Joel Saltz, Emory University — Follow

Included in Health Information Technology Commons, Health Services Research Commons

SHARE

Altmetric 6

Abstract

There is an increasing amount of clinical data in operational electronic health record (EHR) systems. Such data provide substantial opportunities for their re-use for many purposes, including comparative effectiveness research (CER). In a previous paper, we

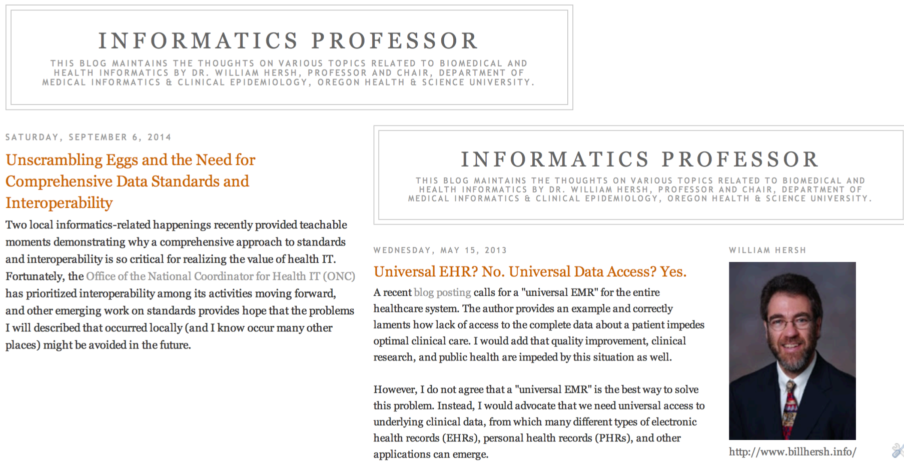
17

Approach: adapt rules of evidence-based medicine (EBM)?

- Ask an answerable question
  - Can question be answered by the data we have?
- Find the best evidence
  - In this case, the best evidence is the EHR data needed to answer the question
- Critically appraise the evidence
  - Does the data answer our question? Are there confounders?
- Apply it to the patient situation
  - Can the data be applied to this setting?

OREGON HEALTH & SCIENCE UNIVERSITY

18



INFORMATICS PROFESSOR

THIS BLOG MAINTAINS THE THOUGHTS ON VARIOUS TOPICS RELATED TO BIOMEDICAL AND HEALTH INFORMATICS BY DR. WILLIAM HERSH, PROFESSOR AND CHAIR, DEPARTMENT OF MEDICAL INFORMATICS & CLINICAL EPIDEMIOLOGY, OREGON HEALTH & SCIENCE UNIVERSITY.

SATURDAY, SEPTEMBER 6, 2014

Unscrambling Eggs and the Need for Comprehensive Data Standards and Interoperability

Two local informatics-related happenings recently provided teachable moments demonstrating why a comprehensive approach to standards and interoperability is so critical for realizing the value of health IT. Fortunately, the Office of the National Coordinator for Health IT (ONC) has prioritized interoperability among its activities moving forward, and other emerging work on standards provides hope that the problems I will described that occurred locally (and I know occur many other places) might be avoided in the future.

INFORMATICS PROFESSOR

THIS BLOG MAINTAINS THE THOUGHTS ON VARIOUS TOPICS RELATED TO BIOMEDICAL AND HEALTH INFORMATICS BY DR. WILLIAM HERSH, PROFESSOR AND CHAIR, DEPARTMENT OF MEDICAL INFORMATICS & CLINICAL EPIDEMIOLOGY, OREGON HEALTH & SCIENCE UNIVERSITY.

WEDNESDAY, MAY 15, 2013

Universal EHR? No. Universal Data Access? Yes.

A recent blog posting calls for a "universal EMR" for the entire healthcare system. The author provides an example and correctly laments how lack of access to the complete data about a patient impedes optimal clinical care. I would add that quality improvement, clinical research, and public health are impeded by this situation as well.

However, I do not agree that a "universal EMR" is the best way to solve this problem. Instead, I would advocate that we need universal access to underlying clinical data, from which many different types of electronic health records (EHRs), personal health records (PHRs), and other applications can emerge.

WILLIAM HERSH

http://www.billhersh.info/

Need for standards and interoperability

- Recognition by ONC as critical for HIT success
- Emerging standards should facilitate
  - e.g., Fast Health Interoperability Resources (FHIR)
    - http://wiki.hl7.org/index.php?title=FHIR_for_Clinical_Users

19

Also need academic contributions of informatics

- Informatics workforce and its training (Hersh, 2010)
  – Development and implementation driven with users and optimal uses in mind – engage by providing value
  – Led by well-trained workforce, including clinical informatics subspecialists (Detmer, 2014)
- Research agenda – must better understand
  – Biases healthcare process creates in EHR data
  – Workflows – impact and optimization
  – User interfaces that allow the entry of high-quality data in time-efficient manner
  – Limitations of all data and how it can be improved

20

OREGON HEALTH & SCIENCE UNIVERSITY

An instance of research: cohort discovery

- Adapting information retrieval techniques to medical records
- Use case somewhat different from usual information retrieval: want to retrieve records and data within them to identify patients who might be candidates for clinical studies
- Another goal: working with large quantity of data, i.e., not few hundred documents typical to natural language processing studies

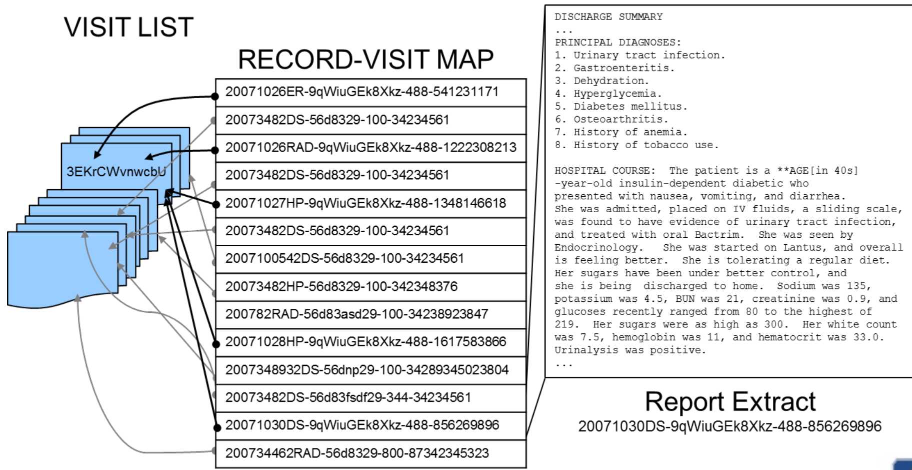
21

Challenges for informatics research with medical records

- Has always been easier with knowledge-based content than patient-specific data due to a variety of reasons
  - Privacy issues
  - Task issues
- Facilitated with development of large-scale, de-identified data set from University of Pittsburgh Medical Center (UPMC)
- TREC Medical Records Track launched in 2011, repeated in 2012 (Voorhees, 2012)

22

Test collection



17,265 visits

101,712 reports (93,552 mapped to visits)

(Courtesy, Ellen Voorhees, NIST)

Some issues for test collection

- De-identified to remove protected health information (PHI), e.g., age number → range
- De-identification precludes linkage of same patient across different visits (encounters)
- UPMC only authorized use for TREC 2011 and TREC 2012 but nothing else, including any other research (unless approved by UPMC)

Easy and hard topics

- Easiest – best median bpref
  - 105: Patients with dementia
  - 132: Patients admitted for surgery of the cervical spine for fusion or discectomy
- Hardest – worst best bpref and worst median bpref
  - 108: Patients treated for vascular claudication surgically
  - 124: Patients who present to the hospital with episodes of acute loss of vision secondary to glaucoma
- Large differences between best and median bpref
  - 125: Patients co-infected with Hepatitis C and HIV
  - 103: Hospitalized patients treated for methicillin-resistant Staphylococcus aureus (MRSA) endocarditis
  - 111: Patients with chronic back pain who receive an intraspinal pain-medicine pump

25

OREGON HEALTH & SCIENCE UNIVERSITY

Failure analysis for 2011 topics (Edinger, 2012)

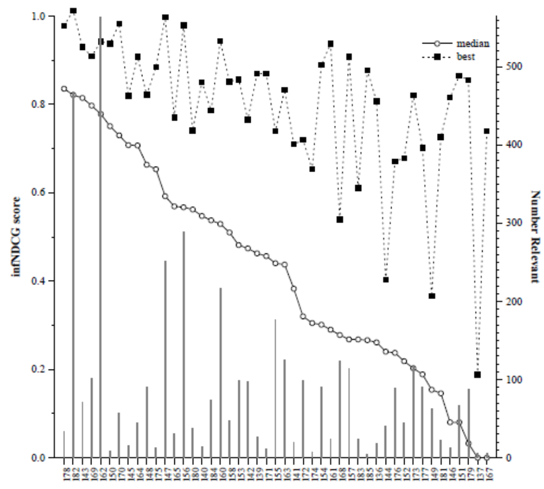
Reasons for Incorrect Retrieval	Number of Visits	Number of Topics
Visits Judged Not Relevant		
Topic terms mentioned as future possibility	16	9
Topic symptom/condition/procedure done in the past	22	9
All topic criteria present but not in the time/sequence specified by the topic description	19	6
Most, but not all, required topic criteria present	17	8
Topic terms denied or ruled out	19	10
Notes contain very similar term confused with topic term	13	11
Non-relevant reference in record to topic terms	37	18
Topic terms not present—unclear why record was ranked highly	14	8
Topic present—record is relevant—disagree with expert judgment	25	11
Visits Judged Relevant		
Topic not present—record is not relevant—disagree with expert judgment	44	21
Topic present in record but overlooked in search	103	27
Visit notes used a synonym or lexical variant for topic terms	22	10
Topic terms not named in notes and must be inferred	3	2
Topic terms present in diagnosis list but not visit notes	5	5

26

OREGON HEALTH & SCIENCE UNIVERSITY

Results for 2012

Run	infNDCG	infAP	P(10)
NLMManual*	0.680	0.366	0.749
udelSUM	0.578	0.286	0.592
sennamed2	0.547	0.275	0.557
ohsuManBool*	0.526	0.250	0.611
atigeo1	0.524	0.224	0.519
UDinfoMed123	0.517	0.236	0.528
uogTrMConQRd	0.509	0.231	0.553
NICTAUBC4	0.487	0.216	0.517



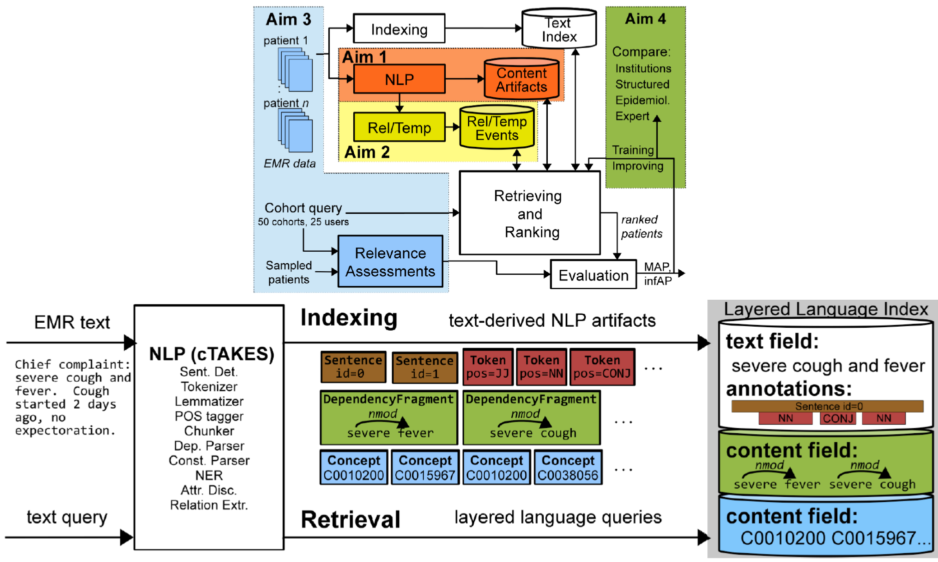
Per-topic scores for infNDCG computed over all runs

What approaches did (and did not) work?

- Best results in 2011 and 2012 obtained from NLM group (Demner-Fushman, 2011)
  - Top results from manually constructed queries using Essie domain-specific search engine (Ide, 2007)
  - Other automated processes fared less well, e.g., creation of PICO frames, negation, term expansion, etc.
- Best automated results in 2011 obtained by Cengage (King, 2011)
  - Filtered by age, race, gender, admission status; terms expanded by UMLS Metathesaurus
- Benefits of approaches commonly successful in IR provided small or inconsistent value
  - Document focusing, term expansion, etc.

OREGON HEALTH & SCIENCE UNIVERSITY

Future work – collaboration with Mayo Clinic; R01 to be funded 9/2014)



Future directions

- Evaluation must focus on real-world
  - Use cases
  - Collections and topics
- Use cases should focus on tasks of clinicians, researchers, and other specific roles
- Collections should reflect type and quantity of information appropriate to use cases

OREGON HEALTH & SCIENCE UNIVERSITY

Conclusions

- There are plentiful opportunities for secondary use or re-use of clinical data
- We must be cognizant of caveats of using operational clinical data
- We must implement best practices for using such data
- We need a research agenda to optimize use

OREGON HEALTH & SCIENCE UNIVERSITY