The Data-Rich, Information-Driven Future of Healthcare

William Hersh, MD
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

Adams, J and Klein, J (2011). Business Intelligence and Analytics in Health Care - A Primer. Washington, DC, The Advisory Board Company. http://www.advisory.com/Research/IT-Strategy-Council/Research-Notes/2011/Business-Intelligence-and-Analytics-in-Health-Care

Berwick, DM and Hackbarth, AD (2012). Eliminating waste in US health care. *Journal of the American Medical Association*. 307: 1513-1516.

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.

Bourgeois, FC, Olson, KL, et al. (2010). Patients treated at multiple acute health care facilities: quantifying information fragmentation. *Archives of Internal Medicine*. 170: 1989-1995.

Brennan, L, Watson, M, et al. (2012). The importance of knowing context of hospital episode statistics when reconfiguring the NHS. *British Medical Journal*. 344: e2432. http://www.bmj.com/content/344/bmj.e2432

Buntin, MB, Burke, MF, et al. (2011). The benefits of health information technology: a review of the recent literature shows predominantly positive results. *Health Affairs*. 30: 464-471.

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.

Chaudhry, B, Wang, J, et al. (2006). Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine*. 144: 742-752.

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.

D'Amore, JD, Sittig, DF, et al. (2012). How the continuity of care document can advance medical research and public health. *American Journal of Public Health*. 102: e1-e4.

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.
Davenport, TH and Harris, JG (2007). Competing on Analytics: The New Science of Winning. Cambridge, MA, Harvard Business School Press.
Davenport, TH, Harris, JG, et al. (2010). Analytics at Work: Smarter Decisions, Better Results. Cambridge, MA, Harvard Business Review Press.
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.
Denny, JC (2012). Mining Electronic Health Records in the Genomics Era. PLOS Computational Biology: Translational Bioinformatics. M. Kann and F. Lewitter.
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.
Diamond, CC, Mostashari, F, et al. (2009). Collecting and sharing data for population health: a new paradigm. *Health Affairs*. 28: 454-466.
Dorr, D, Bonner, LM, et al. (2007). Informatics systems to promote improved care for chronic illness: a literature review. *Journal of the American Medical Informatics Association*. 14: 156-163.
Dorr, DA, Wilcox, AB, et al. (2008). The effect of technology-supported, multidisease care management on the mortality and hospitalization of seniors. *Journal of the American Geriatrics Society*. 56: 2195-2202.
Finnell, JT, Overhage, JM, et al. (2011). All health care is not local: an evaluation of the distribution of emergency department care delivered in Indiana. *AMIA Annual Symposium Proceedings*, Washington, DC. 409-416
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.
Goldzweig, CL, Towfigh, A, et al. (2009). Costs and benefits of health information technology: new trends from the literature. *Health Affairs*. 28: w282-w293.
Greene, SM, Reid, RJ, et al. (2012). Implementing the learning health system: from concept to action. *Annals of Internal Medicine*. 157: 207-210.
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 (2004). Health care information technology: progress and barriers. *Journal of the American Medical Association*. 292: 2273-2274.

Hersh, W (2005). Evaluation of biomedical text mining systems: lessons learned from information retrieval. *Briefings in Bioinformatics*. 6: 344-356.
Hersh, WR, Weiner, MG, et al. (2013). Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical Care*: In press.
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
Hripcsak, G and Albers, DJ (2012). Next-generation phenotyping of electronic health records. *Journal of the American Medical Informatics Association*. 20: 117-121.
Hripcsak, G, Friedman, C, et al. (1995). Unlocking clinical data from narrative reports: a study of natural language processing. *Annals of Internal Medicine*. 122: 681-688.
Hsiao, CJ and Hing, E (2012). Use and Characteristics of Electronic Health Record Systems Among Office-based Physician Practices: United States, 2001-2012. Atlanta, GA, National Center for Health Statistics, Centers for Disease Control and Prevention. http://www.cdc.gov/nchs/data/databriefs/db111.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.
Kern, LM, Malhotra, S, et al. (2013). Accuracy of electronically reported "meaningful use" clinical quality measures: a cross-sectional study. *Annals of Internal Medicine*. 158: 77-83.
King, J, Patel, V, et al. (2012). Physician Adoption of Electronic Health Record Technology to Meet Meaningful Use Objectives: 2009-2012. Washington, DC, Office of the National Coordinator for Health Information Technology. http://www.healthit.gov/sites/default/files/onc-data-brief-7-december-2012.pdf
Kosinski, M, Stillwell, D, et al. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*: Epub ahead of print.
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.
Lee, SJ and Walter, LC (2011). Quality indicators for older adults: preventing unintended harms. *Journal of the American Medical Association*. 306: 1481-1482.
Lewis, M (2004). Moneyball: The Art of Winning an Unfair Game. New York, NY, W. W. Norton & Company.
Longworth, DL (2011). Accountable care organizations, the patient-centered medical home, and health care reform: what does it all mean? *Cleveland Clinic Journal of Medicine*. 78: 571-589.
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, K (2011). Big Data Analytics in Biomedical Research. Biomedical Computation Review. http://biomedicalcomputationreview.org/sites/default/files/w12_f1-big_data.pdf
Mirnezami, R, Nicholson, J, et al. (2012). Preparing for precision medicine. *New England Journal of Medicine*. 366: 489-491.
Nadkarni, PM, Ohno-Machado, L, et al. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*. 18: 544-551.

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.
Rhoads, J and Ferrara, L (2012). Transforming Health Care Through Better Use of Data. Falls Church, VA, Computer Sciences Corp. http://www.csc.com/health_services/insights/80497-transforming_healthcare_through_better_use_of_data
Roebuck, C (2012). The importance of knowing context of hospital episode statistics when reconfiguring the NHS. *British Medical Journal*: Rapid Response. http://www.bmj.com/content/344/bmj.e2432/rr/578977
Ruiz, EJ, Hristidis, V, et al. (2012). Correlating financial tme series with micro-blogging activity. *Fifth ACM International Conference on Web Search & Data Mining*, Seattle, WA. http://www.cs.ucr.edu/~vagelis/publications/wsdm2012-microblog-financial.pdf
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.
Salant, JD and Curtis, L (2012). Nate Silver-Led Statistics Men Crush Pundits in Election. Bloomberg, November 7, 2012. http://www.bloomberg.com/news/2012-11-07/nate-silver-led-statistics-men-crush-pundits-in-election.html
Salathe, M, Vu, DQ, et al. (2013). The dynamics of health behavior sentiments on a large online social network. *EPJ Data Science*. 2: 4. http://www.epjdatascience.com/content/2/1/4/
Savitz, L, Bayley, KB, et al. (2012). Challenges in using electronic health record data for CER: experience of four learning organizations. *Journal of the American Medical Informatics Association*: In review.
Scherer, M (2012). Inside the Secret World of the Data Crunchers Who Helped Obama Win. Time, November 7, 2012. http://swampland.time.com/2012/11/07/inside-the-secret-world-of-quants-and-data-crunchers-who-helped-obama-win/
Schoen, C, Osborn, R, et al. (2012). A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. *Health Affairs*. 31: 2805-2816.
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
Smith, M, Saunders, R, et al. (2012). Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC, National Academies Press.
Stanfill, MH, Williams, M, et al. (2010). A systematic literature review of automated clinical coding and classification systems. *Journal of the American Medical Informatics Association*. 17: 646-651.
Stecker, EC (2013). The Oregon ACO experiment — bold design, challenging execution. *New England Journal of Medicine*. 368: 982-985.
Steffes, T (2012). Predictive analytics: saving lives and lowering medical bills. Analytics, January/February 2012. http://www.analytics-magazine.org/januaryfebruary-2012/505-predictive-analytics-saving-lives-and-lowering-medical-bills

Sun, J, Hu, J, et al. (2012). Combining knowledge and data driven insights for identifying risk factors using electronic health records. *AMIA 2012 Annual Symposium*, Chicago, IL. 901-910
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.
Ugander, J, Karrer, B, et al. (2011). The Anatomy of the Facebook Social Graph. Palo Alto, CA, Facebook. http://arxiv.org/abs/1111.4503
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.
Zhang, Z and Sun, J (2010). Interval censoring. *Statistical Methods in Medical Research*. 19: 53-70.

The Data-Rich, Information-Driven Future of Healthcare

William Hersh, MD
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

OREGON HEALTH & SCIENCE UNIVERSITY

Outline

- Our dysfunctional healthcare system and a vision for fixing and optimizing it
- Part of the solution includes adoption of electronic health records (EHRs) and other technologies
- Toward the data-rich, information-driven learning health system
- Challenges in getting to such a system
- The skills and workforce need to get there

OREGON HEALTH & SCIENCE UNIVERSITY

Our healthcare system is broken in many ways and needs fixin'

- Action must be taken to address (Smith, 2012)
  - $750B in waste (out of $2.5T system)
  - 75,000 premature deaths
- Sources of waste – from Berwick (2012)
  - Unnecessary services provided
  - Services inefficiently delivered
  - Prices too high relative to costs
  - Excess administrative costs
  - Missed opportunities for prevention
  - Fraud
- One vision for repair is the IOM's "learning healthcare system" (Smith, 2012)

http://www.iom.edu/Reports/2012/Best-Care-at-Lower-Cost-The-Path-to-Continuously-Learning-Health-Care-in-America.aspx

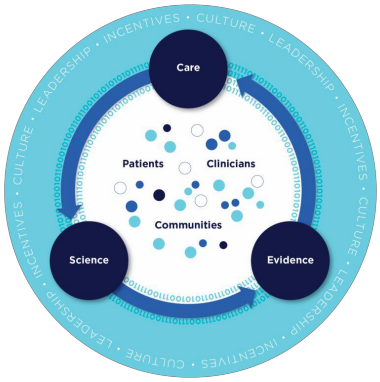
3

OREGON HEALTH & SCIENCE UNIVERSITY

From:

(Smith, 2012)

To:



4

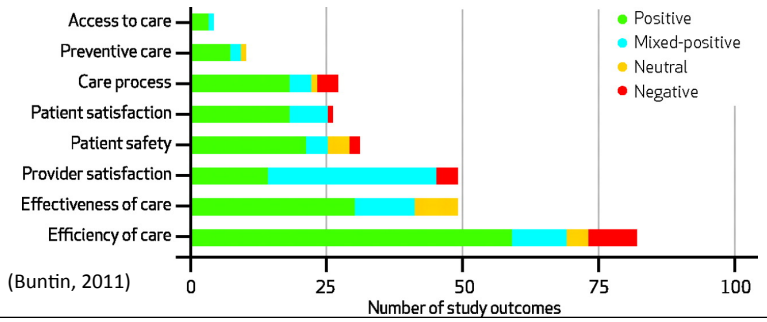
Recommendations for *Best Care, Lower Cost* (Smith, 2012)

- Foundational elements
  - Digital infrastructure
  - Data utility
- Care improvement targets
  - Clinical decision support
  - Patient-centered care
  - Community links
  - Care continuity
  - Optimized operations
- Supportive policy environment
  - Financial incentives
  - Performance transparency
  - Broad leadership

OREGON HEALTH & SCIENCE UNIVERSITY

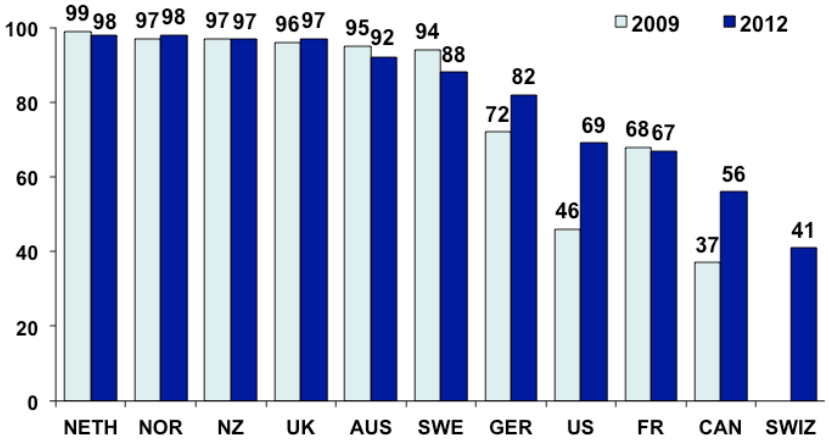
Health information technology (HIT) is part of solution

- Systematic reviews (Chaudhry, 2006; Goldzweig, 2009; Buntin, 2011) have identified benefits in a variety of areas
  - Although 18-25% of studies come from a small number of 'health IT leader" institutions



(Buntin, 2011)

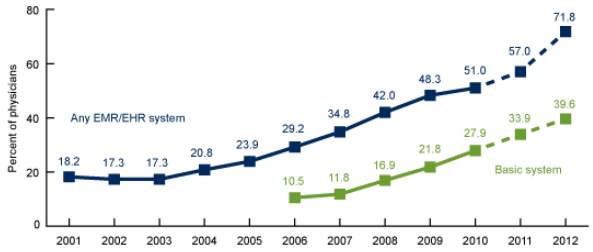
The world is adopting EHRs (Schoen, 2012)



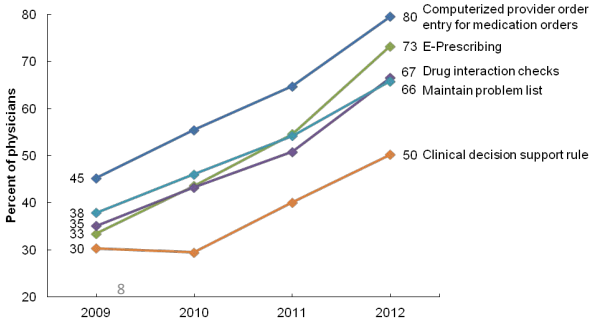
OREGON HEALTH & SCIENCE UNIVERSITY

Even in the US

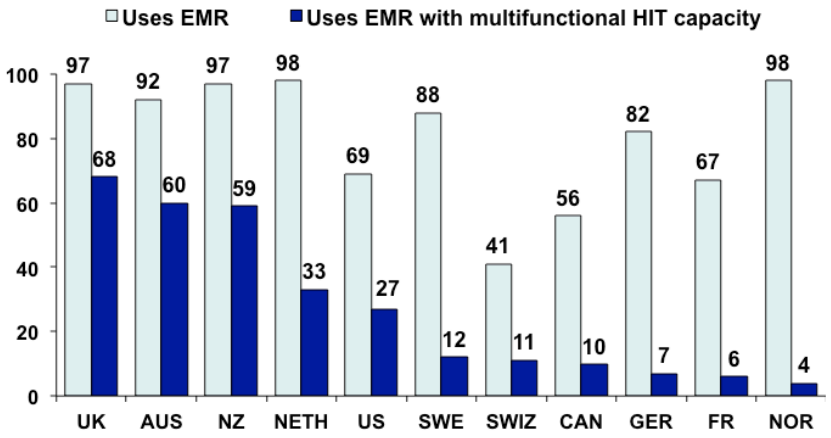
(Hsaio, CDC, 2012)



(King, ONC, 2012)



Although advanced functionality is less common (Schoen, 2012)



Multifunctional health IT capacity – use of at least two electronic functions: order entry management, generating patient information, generating panel information, and routine clinical decision support

OREGON HEALTH & SCIENCE UNIVERSITY

Why has it been so difficult to get there? (Hersh, 2004)

Health Care Information Technology
Progress and Barriers

William Hersh, MD

IN THE 3 DECADES SINCE THE TERM "MEDICAL INFORMATICS" was first used, individuals working at the intersection of information technology (IT) and medicine have developed and evaluated computer applications aim-

in this issue of JAMA, Slack demonstrates the value that patient-physician e-mail can have in improving patient care, and also catalogs the incomplete but encouraging underlying evidence.[11] As with many applications of IT, the technology can improve the existing situation but also empower clinicians and patients to think more fundamentally about how innovation can lead to changes in the way medicine is practiced

- Cost
- Technical challenges
- Interoperability
- Privacy and confidentiality
- Workforce

care IT.[8,9] It is no exaggeration to declare that the years ahead portend the "decade of health information technology."[10]

Informatics is poised to have a major impact in patient-clinician communication. In the Clinical Crossroads article

ment. The rest goes to those who typically do not pay for

Author Affiliation: Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland.
Corresponding Author: William Hersh, MD, Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University School of Medicine, 3181 SW Sam Jackson Park Rd, BICC, Portland, OR 97201-3098 (hersh@ohsu.edu).

See also p 2255.

(Reprinted) JAMA, November 10, 2004—Vol 292, No. 18 2273

OREGON HEALTH & SCIENCE UNIVERSITY

But now we have substantial investment in 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

Setting the stage for a data-rich, information-driven healthcare future

- Use of analytics and/or business intelligence (BI)
  - Analytics is the use of data collection and analysis to optimize decision-making (Davenport, 2010)
  - BI is the "processes and technologies used to obtain timely, valuable insights into business and clinical data" (Adams, 2011)
- As in many areas of advanced information and technology, healthcare and biomedicine are behind the curve of other industries (Miller, 2011; Rhoads, 2012)

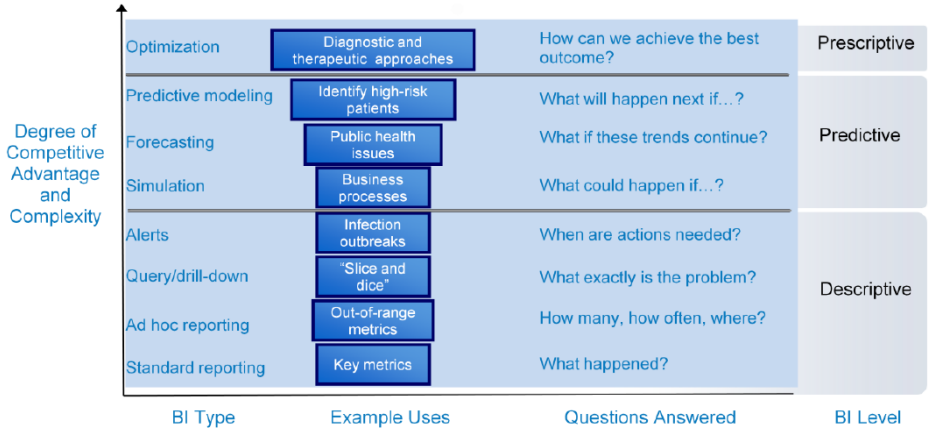
OREGON HEALTH & SCIENCE UNIVERSITY

Who employs analytics and BI outside of healthcare?

- Amazon and Netflix recommend books and movies with great precision
- Many sports teams, such as the Oakland Athletics and New England Patriots, have used "moneyball" to select players, plays, strategies, etc. (Lewis, 2004; Davenport, 2007)
- Facebook can target advertising (Ugander, 2011) and predict your location, political views, sexual preferences, and intelligence (Kosinski, 2013)
- Twitter volume and other linkages can predict stock market prices (Ruiz, 2012), though may be more effective at spreading misinformation (Salathe, 2013)
- Recent US election showed value of using data: re-election of President Obama (Scherer, 2012) and predictive ability of Nate Silver (Salant, 2012)

OREGON HEALTH & SCIENCE UNIVERSITY

Levels of BI (Adams, 2011)



OREGON HEALTH & SCIENCE UNIVERSITY

Analytics part of larger "secondary use" or "re-use" of clinical data

- Many secondary uses or re-uses of electronic health record (EHR) data (Safran, 2007); these include
  - 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

15

OREGON HEALTH & SCIENCE UNIVERSITY

Using data analytics 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)
- Being applied by many now to problem of early hospital re-admission, in particular within 30 days (Sun, 2012; Gildersleeve, 2013)
- Prediction not only of patient response but also behavior, e.g., regimen adherence (Steffes, 2012)
- A requirement of coming "precision medicine" (Mirnezami, 2012)

16

OREGON HEALTH & SCIENCE UNIVERSITY

Quality measurement and improvement

- Quality measures increasingly used in US and elsewhere
- Use has been more for process than outcome measures (Lee, 2011), e.g., Stage 1 meaningful use

NQF Measure Number & PQRI Implementation Number	Clinical Quality Measure Title
Core Clinical Quality Measures	
NQF 0013	Hypertension: Blood Pressure Measurement
NQF 0028	Preventive Care and Screening Measure Pair: a) Tobacco Use Assessment, b) Tobacco Cessation Intervention
NQF 0421 PQRI 128	Adult Weight Screening and Follow-up
Alternate Clinical Quality Measures	
NQF 0024	Weight Assessment and Counseling for Children and Adolescents
NQF0041 PQRI 110	Preventive Care and Screening: Influenza Immunization for Patients 50 Years Old or Older
NQF 0038	Childhood Immunization Status

OREGON HEALTH & SCIENCE UNIVERSITY

Quality measurement and improvement

- 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; Kern, 2013)

OREGON HEALTH & SCIENCE UNIVERSITY

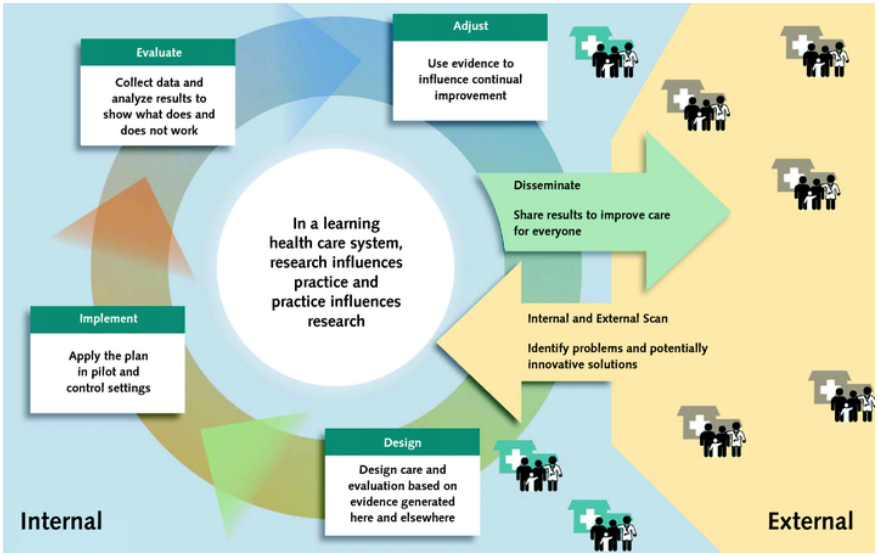
Clinical and translational research

- Led in part by activities of NIH Clinical and Translational Science Award (CTSA) Program (Mackenzie, 2012)
- eMERGE Network – connecting genotype-phenotype, 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), while blood cell count abnormalities (Crosslin, 2012), thyroid disorders (Denny, 2011), etc.
- Other successes include replication of RCTs findings of
  - 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)

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)
- One notable success is Google Flu Trends (http://www.google.org/flutrends/) – search terms entered into Google predicted flu activity but not enough to intervene (Ginsberg, 2009); last year, performance was poorer (Butler, 2013)

Implementing the learning healthcare system (Greene, 2012)



Caveats for use of operational EHR data (Hersh, 2013) – may be

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

OREGON HEALTH & SCIENCE UNIVERSITY

Inaccurate

- Documentation not always a top priority for busy clinicians (de Lusignan, 2005)
- Data entry errors in a recent analysis in the English National Health Service (NHS) – yearly hospital statistics showed approximately (Brennan, 2012)
  - 20,000 adults attending pediatric outpatient services
  - 17,000 males admitted to obstetrical inpatient services – mainly due to male newborns (Roebuck, 2012)
  - 8,000 males admitted to gynecology inpatient services

Inaccurate (cont.)

- 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 (Savitz, 2012)
  - Missing
  - Erroneous
  - Un-interpretable
  - Inconsistent
  - Inaccessible in text notes

24

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)
- Quality measures under-reported based on under-capture of data due to variation in clinical workflow and documentation practices (Parsons, 2012)
- Ability to identify diabetic patients successively increased as time frame of assessing records was increased from one through ten years of analysis (Wei, 2013)

OREGON HEALTH & SCIENCE UNIVERSITY

Incomplete (cont.)

- Studies of health information exchange (HIE)
  - Study of 3.7 million patients in Massachusetts found 31% visited two or more hospitals over five years (57% of all visits) and 1% visited five or more hospitals (10% of all visits) (Bourgeois, 2010)
  - Analysis of 2.8 million emergency department patients in Indiana found 40% had data at multiple institutions (Finnell, 2011)

OREGON HEALTH & SCIENCE UNIVERSITY

Unrecoverable for research

- Many clinical data are "locked" in narrative text reports (Hripcsak, 1995; Hripcsak, 2012), including summaries of care (D'Amore, 2012)
- A promising approach for recovering these data for research is natural language processing (NLP) (Nadkarni, 2011)
  - Has been most successful when applied to the determination of specific data elements, e.g., eMERGE studies (Denny, 2012)
- State of the art for performance of NLP has improved dramatically over the last couple decades, but is still far from perfect (Stanfill, 2010)
  - Still do not not know how good is "good enough" for NLP in data re-use for research, quality, etc. (Hersh, 2005)

Many data "idiosyncrasies"

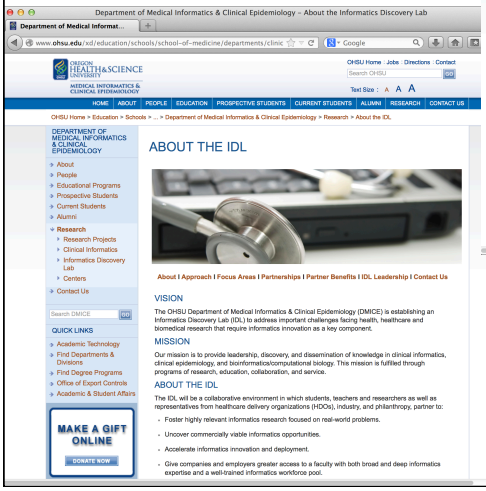
- "Left censoring": First instance of disease in record may not be when first manifested
- "Right censoring": Data source may not cover long enough time interval
- Data might not be captured from other clinical (other hospitals or health systems) or non-clinical (OTC drugs) settings
- Bias in testing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

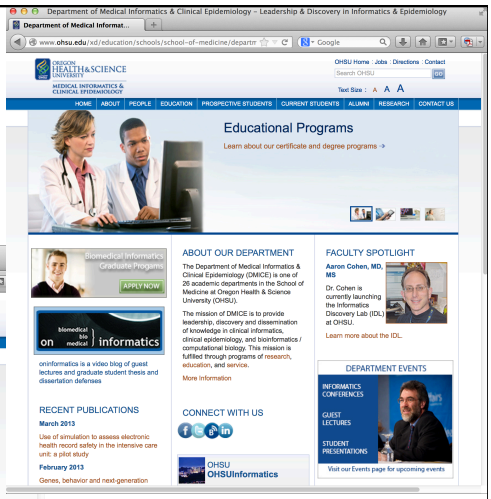
Most important requirement may be changing healthcare system

- US healthcare system still mostly based on fee for service model – little incentive for managing care in coordinated manner
- Informatics tools can help in care coordination (Dorr, 2007; Dorr, 2008)
- Primary care medical home (PCMH) might be first step to improving value and providing incentive for better use of data (Longworth, 2011)
- Affordable Care Act (ACA) implements accountable care organizations (ACOs), which provide bundled, quality-adjusted payments for conditions (Longworth, 2011)
  - Being implemented statewide in Oregon Medicaid system (Stecker, 2013)

OREGON HEALTH & SCIENCE UNIVERSITY

Another need is for academic programs leading in research and education





www.ohsu.edu/informatics

www.ohsu.edu/idl

OREGON HEALTH & SCIENCE UNIVERSITY

DMICE/IDL faculty areas of expertise

Expertise

- **Healthcare analytics**
- **Workflow redesign**
- **Qualitative and quantitative evaluation methodologies**
- **Natural language processing, machine learning, and information retrieval**
- **Biomedical terminologies, ontologies, and coding**
- **Biomedical data structures, representation, and normalization**

Focus Areas

- **Predictive Analytics for Healthcare Process Redesign**
- **Usability for EHR Data and Users**
- **Population Management and Care Coordination Information Systems**
- **Mobile Telemedicine**
- **Precision Medicine**

OREGON HEALTH & SCIENCE UNIVERSITY

Conclusions

- A growing body of evidence supports EHR and other IT to improve health and healthcare
- The world is gradually adopting EHRs and other IT
- The next step is to make use of the increasing data through analytics and BI to achieve the learning healthcare system
- There are challenges, but also benefits, to this use data-driven, information-driven evolution
- There is a growing need for research and education in all areas of informatics

OREGON HEALTH & SCIENCE UNIVERSITY

For more information

- Bill Hersh
  - http://www.billhersh.info
- Informatics Professor blog
  - http://informaticsprofessor.blogspot.com
- OHSU Department of Medical Informatics & Clinical Epidemiology (DMICE)
  - http://www.ohsu.edu/informatics
  - http://www.youtube.com/watch?v=T-74duDDvwU
  - http://oninformatics.com
- What is Biomedical and Health Informatics?
  - http://www.billhersh.info/whatis
- Office of the National Coordinator for Health IT (ONC)
  - http://www.healthit.gov
- American Medical Informatics Association (AMIA)
  - http://www.amia.org
- National Library of Medicine (NLM)
  - http://www.nlm.nih.gov

OREGON HEALTH & SCIENCE UNIVERSITY