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.

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

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.

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.

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.

Denny, JC (2012). Mining Electronic Health Records in the Genomics Era. <u>PLOS Computational Biology: Translational Bioinformatics</u>. 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.

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.

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.

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 (2005). Evaluation of biomedical text mining systems: lessons learned from information retrieval. *Briefings in Bioinformatics*. 6: 344-356.

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

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 (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.

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.

Nadkarni, PM, Ohno-Machado, L, et al. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*. 18: 544-551.

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.

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

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.

Seiler, KP, Bodycombe, NE, et al. (2011). Master data management: getting your house in order. *Combinatorial Chemistry & High Throughput Screening*. 14: 749-756.

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.

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.

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

1

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

2

Outline

- Our dysfunctional healthcare system and information-driven solutions for improving it
- 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

3

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

BEST CARE AT LOWER COST

The Path to Continuously Learning Health Care in America

WHAT'S POSSIBLE FOR HEALTH CARE?

- 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

4

2

From:

(Smith, 2012)

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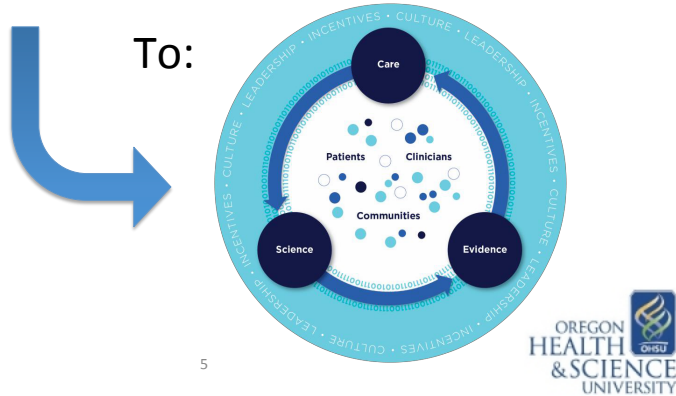
US has made substantial investment in health information technology (HIT)

"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)

Which has led to significant EHR adoption in the US

(Hsiao, 2014)

(Charles, 2014)
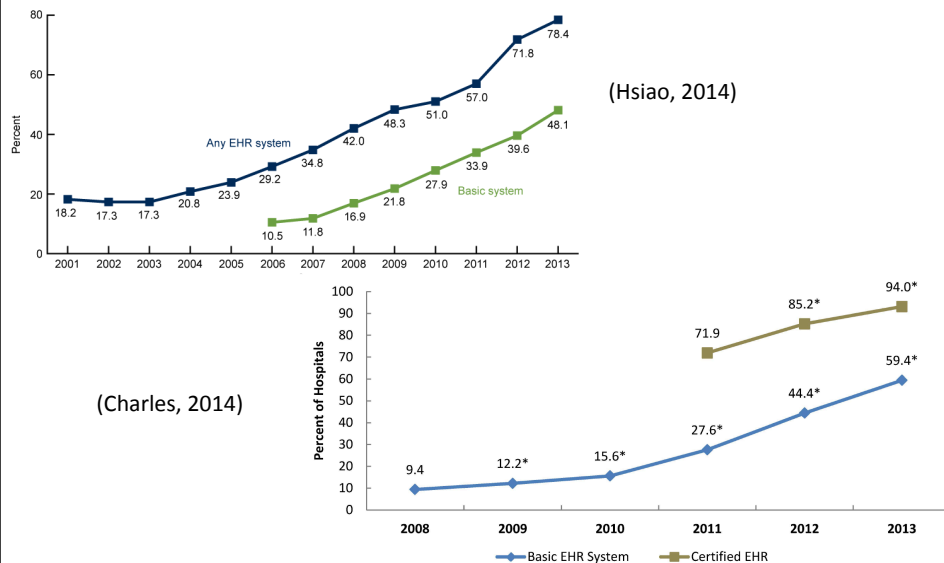
Providing opportunities for "secondary use" or "re-use" of clinical data

- (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

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)

9

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)

10

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.

11

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)

12

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)

OREGON
HEALTH
&SCIENCE
UNIVERSITY

13

Implementing the learning healthcare system (Greene, 2012)
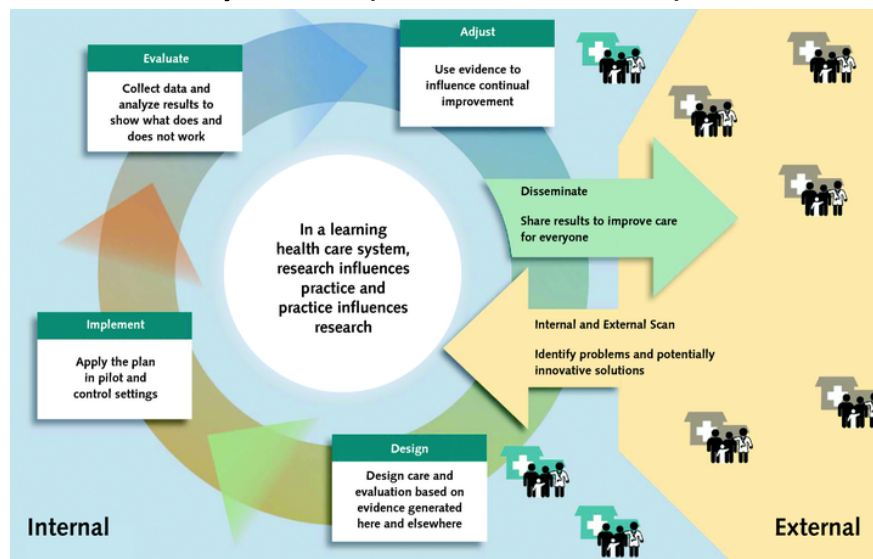
Caveats for the Use of Operational Electronic Health Record Data in Comparative Effectiveness Research

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

15

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

OREGON
HEALTH
&SCIENCE
UNIVERSITY

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 (Bayley, 2013)
  - Missing
  - Erroneous
  - Un-interpretable
  - Inconsistent
  - Inaccessible in text notes

17

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)

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)

19

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)

Of unknown provenance and insufficient granularity

- Provenance – knowing where your data come from (Seiler, 2011)
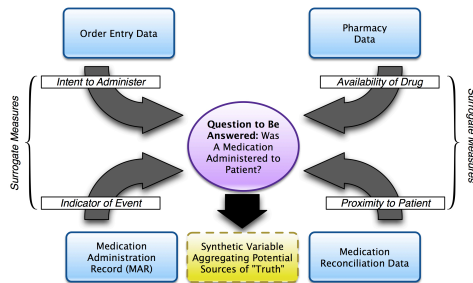
- Granularity – knowing what your data mean
  - Diagnostic codes assigned for billing purposes may be generalized to a broad class of diagnosis due to regulatory and documentation requirements
  - For example, patient with set of complex cytogenetic and morphologic indicators of a pre-leukemic state may be described as having "myelodysplastic syndromes (MDS)" for billing purposes, but this is insufficient for other purposes, including research

21

Many data "idiosyncrasies" between clinical practice and research protocols

- "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

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
- Develop clinical data research networks
- Appropriate use of informatics expertise

23

OREGON
HEALTH
&SCIENCE
UNIVERSITY

Assessing and using data

- Assessing availability, completeness, quality, and transformability
  - Need to address small data issues before getting to big data
- Also need
  - Toolkits and pipelines
  - Methods and metrics for assessment
  - Standardized reporting methods

24

OREGON
HEALTH
&SCIENCE
UNIVERSITY
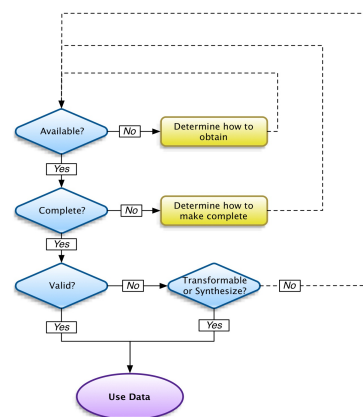
Approach: adapt rules of evidence-based medicine (EBM)?

- 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?

25

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

26

Develop clinical data research networks

- PCORnet developing
  - Clinical data research networks (CDRNs) – 11 networks aggregating data on >1M patients each
    - (Fleurence, 2014; Collins, 2014; and other papers in JAMIA special issue)
  - Patient-powered research networks (PPRNs) – 18 networks of activated patients focusing on specific diseases
  - Common Data Model for all data

http://pcornet.org/
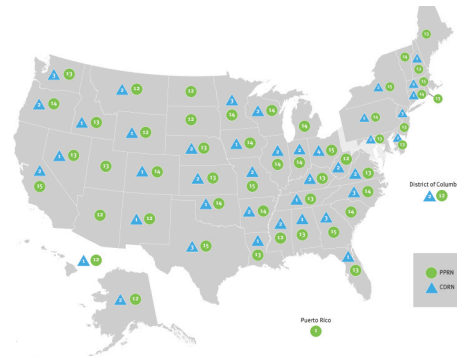
27

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

28

14