

Truth Versus Truthiness in Clinical Data

Mark Weiner, MD, FACP, FACMI

Assistant Dean for Informatics, Temple University School of Medicine

mark.weiner@tuhs.temple.edu



Disclosures

- Relevant Financial Disclosures
 - No Relevant Financial Interests
- Non-FDA Approved uses
 - No Non-FDA Approved Uses

Mark G. Weiner, MD

My mantra

In attempting to arrive at the truth, I have applied everywhere for information, but in scarcely an instance have I been able to obtain hospital records fit for any purpose of comparison. If they could be obtained, they would enable us to decide many other questions besides the one alluded to. They would show the subscribers how their money was being spent, what good was really being done with it, or whether the money was not doing mischief rather than good.

Truth

Definition by Merriam-Webster

(1) : the state of being the case : fact

(2) : the body of real things, events, and facts : actuality

Truthiness

- Merriam-Webster's **Word of the Year for 2006**:
- the quality of preferring concepts or facts ***one wishes to be true***, rather than concepts or facts ***known to be true***
- How do we **know** something is **true**, especially in medicine?

Capturing the Truth

Head versus Heart

- Are men being overtreated, or are women being undertreated for heart disease?
 - Men do get more invasive cardiac procedures than women, but...
- Should all post menopausal women use hormone replacement therapy?
 - We used to think so, but...
- Should all women between 40-50 receive mammograms?
 - Some women get cancer below age 50, but...
- Should medical care be a 24/7 operation?
 - Seems to make sense, but..



Informatics and Data Analytics

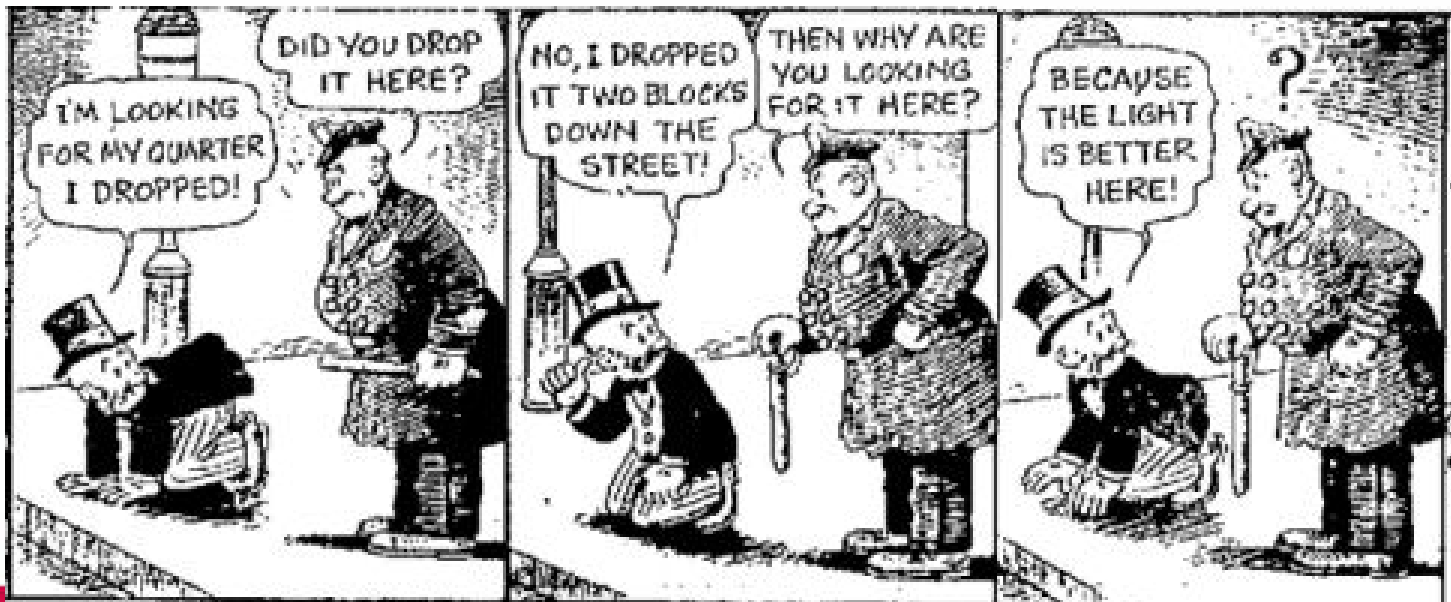
Data and Analysis

- How do good informatics methods lead to the truth?
- How many ways can informatics methods lead you astray of the truth?
- What assumptions do you make about the validity of the data being processed?
- What assumptions do you make about the appropriate analysis?
- Have your assumptions in transforming raw data to an analytical data set introduced bias?

Informatics and Data Analytics

Data and Analysis

- What assumptions do you make about the appropriateness of the data for the truth you seek?



Informatics and Data Analytics

Interpretation and implementation of results

- What assumptions do you make about the interpretation of the results?
 - Statistical significance versus clinical significance
- Do your results advance/alter conclusions made through other means?
- Is there a threshold where your results change clinical decisions
- How do you KNOW your results are correct?

Data Substrate: Informatics and “Big” Data Analytics

- High Volume/Velocity/Variety/Veracity/Validity/Volatility, Visualization/Value of data that is sensibly related to the clinical issue of interest
- Bigger Not Necessarily Better
 - Big data is not necessarily representative data
 - Will participants in the Precision Medicine Initiative reflect patients in North Philadelphia?
 - Do you have high confidence in the gold standard for exposure and outcome?
 - Do the findings support a clinical decision?

My type of Data Sources

Not necessarily big data

- Clinical and administrative data collected through the routine operation of the health system
 - Visits and missed visits
 - vital signs
 - Lab/ancillary study (PFTs, Echocardiograms, radiology) results
 - meds prescribed/picked-up/taken
 - diagnoses assigned
 - procedures performed
 - NLP-requiring things like clinic notes, and other narrative reports
 - Social History
 - Telemetry monitoring

My type of Data Sources

- Sparse
- Not continuous
- Systematically and non-systematically missing data
- Temporal Pattern of data, regardless of the results, MAY (but not necessarily) be associated with the outcome
- Disagreement on the presence, absence, or even the nature of the outcome
- Have not seen machine learning algorithms that account for all this variability in raw clinical data.

Common assumptions that are not quite true

Data Accuracy

- **The billing data is so inaccurate, it cannot be used for research**
 - If you alter your notion of a billing code from “the patient definitely has this disease” to “the patient has something like this disease” or “the doctor is worried about this disease,” and incorporate that ambiguity into your analysis, you will be much happier.
 - **Embrace the ambiguity and heterogeneity in billing (and all) data!**

Common (data analytics vendor) assumptions that are not quite true

Study feasibility

- If I look retrospectively at observational data and find X patients with an acute disease Y over time period Z, I can expect to enroll X patients with disease Y over time period Z
 - Your ability to retrospectively recognize patients who have a disease does NOT mean you can proactively identify them with enough lead time to enroll them in a trial (eg interventions to prevent ICU admissions)
- I need to program my database query tool with ALL inclusion and exclusion criteria to assess study feasibility
 - Especially with some rare exclusion criteria, sometimes it is better to “play the odds,” and not actually query on every criteria

Common assumptions that are not quite true

Reporting Results

- **Results of observational analyses should be presented in a manner analogous to clinical trials** – the relationship between an exposure and outcome should be expressed as a single point estimate and confidence interval
 - AT BEST – ignores the robustness of similar or alternate findings generated under different assumptions
 - Your one answer may be spuriously right or wrong.
 - AT WORST – a finding may be a cherry-picked result

Common assumptions that are not quite true

Shared Computable Phenotypes will create more uniform cohorts across different institutions

- Incorporating all the ways a disease can be identified to capture local variation in coding style or test ordering may create false equivalencies
 - Diabetes by diagnosis, lab or medication use
 - Cohorts identified by high A1c with no DM diagnoses will tend to have higher A1cs than cohorts identified with diagnoses
 - “At Least 2 Diabetes Diagnoses”
 - Certainly more likely to truly have diabetes, but need to look at distribution of number of diagnoses and the average A1c – More Diabetes Diagnoses correlates to higher A1cs
- Is it OK to have different disease definitions for Research, Quality measurement?

Common assumptions that are not quite true

Tweaking the Computable Phenotype to expand The pool of recruitable patients is OK

- Power analysis says you need 1000 patients at increased cardiac risk
- You only have 800 patients
- You expand ever-so-slightly the age, BP and cholesterol criteria in your computable phenotype to get 1000 patients
- When you run the study, the interim analysis shows a non-significant difference in outcomes between intervention and control because the event rate in the control group was lower than expected.

Pioglitazone and bladder cancer: a propensity score matched cohort study

Li Wei, Thomas M. MacDonald

Medicines Monitoring Unit (MEMO), Division of Clinical Pharmacology, Dundee Medical School, Dundee, UK

BMJ

BMJ 2012;344:e3645 doi: 10.1136/bmj.e3645 (Published 31 May 2012)

Page 1 of 11

BJCP May 2012: "In this study population, pioglitazone does not appear to be significantly associated with an increased risk of bladder cancer in patients with type 2 diabetes."

RESEARCH

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Pioglitazone is mainly used in combination with diet and exercise and other anti-diabetic medications to treat type 2 diabetes mellitus.
- Long term use of pioglitazone (as part of therapy) may be associated with an increased risk of bladder cancer.

The use of pioglitazone and the risk of bladder cancer: a propensity score matched case-control study

BMJ May 2012: "The use of pioglitazone is associated with an increased risk of incident bladder cancer among people with type 2 diabetes."

WHAT THIS STUDY ADDS

- In this study population, pioglitazone does not appear to be significantly associated with an increased risk of bladder cancer in patients with type 2 diabetes.

Jonathan Assayag *graduate student*¹, Agnieszka Majdan *endocrinologist*⁴, Michael N Pollak *oncologist and professor*², Samy Suissa *professor*⁵

Christian B Filion *assistant professor*^{1,3}

¹Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, 3755 Côte Sainte-Catherine, H-425.1, Montreal, Quebec, Canada, H3T 1E2; ²Department of Oncology, McGill University, Montreal, Quebec, Canada; ³Division of Clinical Epidemiology, McGill University, Montreal; ⁴Division of Endocrinology, Jewish General Hospital, Montreal; ⁵Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal

Oral Fluoroquinolones and Risk of Retinal Detachment

Mahyar Eitminan, PharmD, MSc (epi)
 Farzin Foroughian, MD, MSc, FRCSC
 James M. Brophy, MD, PhD, FRCPC
 Steven T. Bird, PharmD
 David Maberley, MD, MSc, FRCSC

Context Fluoroquinolone use is associated with numerous case reports of ocular safety, particularly retinal detachment.
Objective To examine the risk of developing a retinal detachment in patients taking oral fluoroquinolones.
Design, Setting, and Population Retrospective cohort study in British Columbia, Canada.

April 2012: "Patients taking oral fluoroquinolones were at a higher risk of developing a retinal detachment"

ter bu riel. Al all sociated with a wide array of adverse events such as dysglycemia,¹ cardiac arrhythmia,² and neuropsychiatric events.³ Fluoroquinolones also have been linked to several forms of ocular toxicity such as corneal perforations,⁴ optic neuropathy,⁵ and retinal hemorrhages.⁶ In 2011, the label for gemifloxacin was updated to include hemorrhage,⁶ which includes retinal hemorrhage that was reported during postmarketing surveillance. A class-wide warning for fluoroquinolones also has been issued for tendon rupture,⁷ which raises concerns for the effect of these drugs on connective tissue in the eye. Animal studies also provide evidence for retinal degeneration with use

a higher risk of developing retinal detachment. The adjusted rate ratio (ARR), 4.0 vs 0.2% of controls; ARR, 6.1% of controls; ARR, 1.1% of controls. The absolute risk increase was 0.1% per person-years (number needed to harm, 1000). There was no evidence of increased risk of retinal detachment and β -lactam antibiotics (ARR, 0.95 [95% CI, 0.85-1.05]).
Conclusion Patients taking oral fluoroquinolones had a higher risk of developing retinal detachment compared with nonusers. The absolute risk increase was small.
 JAMA. 2012;307(13):1414-1419.

through the destructive effects of these drugs on collagen and other connective tissue.¹¹ Collagen fibers play a key role in the structure and function of the retina.

Original Investigation

Association Between Oral Fluoroquinolone Use and Retinal Detachment

Björn Pasternak, MD, PhD; Henrik Svanström, MSc; Mads Melbye, MD, DrMedSci; Anders Hviid, MSc, DrMedSci

IMPORTANCE A recent study of ophthalmologic patients found a strong association between fluoroquinolone use and retinal detachment. Given the prevalent use of fluoroquinolones, this could, if confirmed in the general population, translate to many excess cases of retinal detachment that are potentially preventable.

al fluoroquinolone use is associated with an increased risk of

OBJECTIVES A nationwide, register-based cohort study in Denmark linked data on participant characteristics, filled prescriptions, and cases of retinal detachment with surgical treatment (scleral buckling, vitrectomy, or pneumatic retinopexy). The cohort included 749 702 episodes of fluoroquinolone use

Dec 2013: "Oral fluoroquinolone use was not associated with increased risk of retinal detachment"

MAIN OUTCOME MEASUREMENTS The main outcome was incident retinal detachment. The main variables were the number of retinal detachments and recent use (day

RESULTS A total of 566 cases of retinal detachment occurred, of which 465 (82%) were rhegmatogenous detachments; 72 in fluoroquinolone users and 494 in control nonusers. The crude incidence rate was 25.3 cases per 100 000 person-years in current users, 18.9 in recent users, 26.8 in past users, and 24.8 in distant users compared with 19.0 in nonusers. Compared with nonuse, fluoroquinolone use was not associated with a significantly increased risk of retinal detachment: the adjusted RRs were 1.29 (95% CI, 0.53 to 3.13) for current use;

- ← Editorial page 2151
- ← JAMA Patient Page 2212
- Supplemental content at jama.com

Red flags

- Show me an investigator that comes to a informatician looking for patients to enroll after a study protocol is written, and I'll show you a study that is doomed to fail!
 - Just as you would not develop a study protocol without a statistician at your side, you need a informatician early on in the process
- Study feasibility requires more than asking the investigator about diagnosis codes. Work hard with investigators to create a thorough computable phenotype for the condition of interest.
 - The query process is iterative and should start with an analyst's presentation of at least a frequency count of common codes for a diagnosis or a discussion of disease characteristics
 - Patients with some diagnoses are better found through lab or medication criteria

Informatics solutions with caveats

Focused disease definitions may increase the accuracy of the data, but make the analytical data set less generalizable

- Requiring an echocardiogram to definitively rule in or rule out a diagnosis of CHF limits your cohort to people who were sick enough to require an echo – even among patients who turn out NOT to have CHF by echo

Informatics solutions with caveats

Integrating more databases offers the promise of filling in gaps in the continuum of care,

- But it also increases the likelihood of finding clinical conflicts in the data for an individual
- If you cannot identify patients across different systems, then there will be double-counting of individual

Semantic interoperability will enable different information technology systems to understand the true meaning of data being sent

- But have you ever seen two DOCTORS agree upon the meaning of what they hear?

Informatics solutions with caveats

Standards will enable computer systems to share a common language to describe clinical concepts

- But the precision inherent in these vocabularies often exceeds the precision of medicine
- ICD-10 is bigger than ICD-9, but is it really better?
 - Many of the “new” terms allow more precision in localizing musculoskeletal issues, not capturing nuance in medical issues

The Universal Identifier problem for identifying individuals across systems will be solved with better algorithms

- But then the state of the science ~~will demand~~ already demands defined links between family members!

Informatics solutions with caveats

Automated business rules will help with managing updates to data

- If an address changes, is it because the old address was wrong or did the patient move?
- What if a gender category changes?
- Even if you are SURE an earlier record is an error, should that record be removed from the warehouse?
 - Risk adjustment – requires true data
 - Evaluation of decision support – probably should use data that was known at contemporaneous time
- Should corrections be made to warehouse, or source systems?

Deriving **truth** from the appropriate application of Research Informatics Methods

- **Understand what you can and cannot expect from observational data**
 - Millions of lives and multiple years of overall data does not mean individuals are followed for a long enough time to assess all possible exposure/outcome associations
 - Studying 1000 patients for 1 year is not equivalent to studying 100 patients for 10 years (even though both are 1000 person-years)
 - Big data does NOT always mean REPRESENTATIVE data
 - Results of data analyses are SUPPOSED to provide a broad set of answers under different assumptions of the definition of exposed population, and the complex nature of drug exposure in the real world.

Deriving **truth** from the appropriate application of Research Informatics Methods

- Better accuracy than humans, or ability to augment human assessment is a great goal, but be mindful of what can be done with the improved knowledge that is gained
 - Love truly precise genomic-guided therapy
 - Hate imperfect risk prediction of diseases that have no adequate treatment or ethically-challenging setting of treatment thresholds
- For you big data people
 - Beware impact of heterogeneity that exists within members of the same cluster
 - Be mindful of a tarnished gold standard in your classification analyses
- Never forget the real people underlying the data you use and who are impacted by your findings



Temple University
Health System

Caveats in Data Analytics to Improve Clinical Care

Is the data Accurate?

Accuracy = truth?

- If the information system says a patient has asthma, then the patient has asthma

Accuracy = true representation of the source data?

- If the information system says a patient has asthma, then the source data for the patient includes a code for asthma or other medications or diagnostic testing results consistent with asthma

Caveats in Data Analytics to Improve Clinical Care

Is the data Accurate?

Clinical data is used in real clinical practice so it must be accurate except...

- Physician patient communication/misunderstandings
- Busy doctors don't code/document everything
- Idiosyncrasies of the clinical setting in which data is collected
- Ambiguity inherent in the practice of medicine
- Code creep-
 - Early codes before diagnosis of gallstones is confirmed may suggest simple abdominal pain
 - Changing billing rules can alter the recording of diagnoses

Adapted and expanded from O'Malley KJ, Cook KF, Price MD, 29 KR, Hurdle JF,

Ashton CA Measuring Diagnoses: ICD Code Accuracy Health Services Research 2005. 40:1620-39.



Temple University
Health System

Caveats in Signal Processing to Improve Clinical Care

Research using billing and even EHR data must account for all of the realities inherent in the underlying data

- **Absence of evidence is not evidence of absence**
 - Just because you don't see evidence of a disease doesn't mean the patient doesn't have the disease
- **Vigilance in exploring for a condition, even if it turns out NOT to be present, has clinical relevance**
 - Providers may look harder for the presence of a condition or a physical finding in some patients more than others
- **Clusters of visits are significant, but not necessarily directly related to the condition being studied**
 - Someone with a lot of visits related to asthma may get a cholesterol drawn sooner than a patient without asthma whose cholesterol management is being done more routinely.

Caveats in Data Analytics to Improve Clinical Care

To find patients with a certain disease, you need to consider all the ways the disease may be represented in diagnosis codes and ancillary test results

- URI/bronchitis/tracheitis/pharyngitis/sinusitis are harder to reliably distinguish that most people would like to think.
- Green sputum is not always bacterial and Yellow sputum is not always viral
- Broad definitions must be considered in the assessment of both risks and outcomes

Caveats in Data Analytics to Improve Clinical Care

Left Censoring

- the first instance of a disease in the data is not necessarily the time the disease first appeared.
- Absence of a disease in the data does not mean the condition does not exist

Right Censoring

- Data must cover a time interval long enough for the exposure to result in an outcome that is captured in the data.
- Does the exposure influence the occurrence of the outcome in 1 year, or 5 years?

Caveats in Data Analytics to Improve Clinical Care

Data from uncaptured clinical settings

- Received Angioplasty at big city institution A. Hospitalized at local community hospital B for GI bleed.

Logical inconsistencies across information systems

- Allergy to Drug Z listed from data source A
- Prescriptions for Drug Z from data source B

Temporal issues

- Presence of murmur AFTER echo report (or cardiology visit)
- Heart failure reported AFTER starting ACE-I

Caveats in Data Analytics to Improve Clinical Care

Uncoded information (“institutional metadata”)

- Patients with certain characteristics are cared for in certain locations of the hospital
 - Systematically sicker patients in certain rooms in the ED, or closer to the nursing station.
 - Which rooms are isolation rooms?

Regional (and personal) variation in documentation styles

Caveats in Data Analytics to Improve Clinical Care

Unmeasured confounders

Treatment bias

- Medications are prescribed because the provider believes the patient needs them
- Sicker patients may systematically receive certain types of medicine more than others

Testing bias

- Patients who test positive for certain tests are more likely to receive additional testing (related or unrelated)

Caveats in Data Analytics to Improve Clinical Care

Unmeasured confounders

Formulary issues

- Insurer rules limiting the use of some meds to patients with certain underlying conditions or who have already failed other meds?
- In Phase 4 “real world” clinical trials, a new drug may not perform as well as in head-to-head clinical trials

Practice variation

- Some meds are favored because of non-clinical issues

OTC Meds

- Aspirin, other NSAIDs, PPIs

Available in Medical Care

SUPPLEMENT

Med Care 2013;51: S30–S37

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.artzog, MD, MS,# James J. Cimino, MD,**
and Joel H. Saltz, MD, PhD††*

Abstract: The growing amount of data in operational electronic health record systems provides unprecedented opportunity for its reuse for many tasks, including comparative effectiveness research. However, there are many caveats to the use of such data. Electronic health record data from clinical settings may be inaccurate, incomplete, transformed in ways that undermine their meaning, unrecoverable for research, of unknown provenance, of insufficient granularity, and incompatible with research protocols. However, the quantity and real-world nature of these data provide impetus for their use, and we develop a list of caveats to inform would-be users of such data as well as provide an informatics roadmap that aims to insure this opportunity to augment comparative effectiveness research can be best leveraged.

Key Words: comparative effectiveness research, electronic health records, clinical data, coded (claims) data, biomedical informatics

(Med Care 2013;00: 000–000)

ability to advance biomedical and health care science and practice through the reuse of clinical data.^{2–4} This investment sets the foundation for a “learning” health care system that facilitates clinical research, quality improvement, and other data-driven efforts to improve health.^{5,6}

At the same time, there has also been substantial federal investment in comparative effectiveness research (CER) that aims to study populations and clinical outcomes of maximal pertinence to real-world clinical practice.⁷ These efforts are facilitated by other investments in research infrastructure, such as the Clinical and Translational Research Award (CTSA) program of the US National Institutes of Health.⁸ Many institutions funded by CTSA awards are developing research data warehouses of data derived from operational systems.⁹ Additional federal investment has been provided by the Office of the National Coordinator for Health Information Technology (ONC) through the Strategic Health IT Advanced Research Projects (SHARP) Program,