Benign Urology in the Big Data World

Stephen K. Van Den Eeden, PhD
Research Scientist III, Division of Research

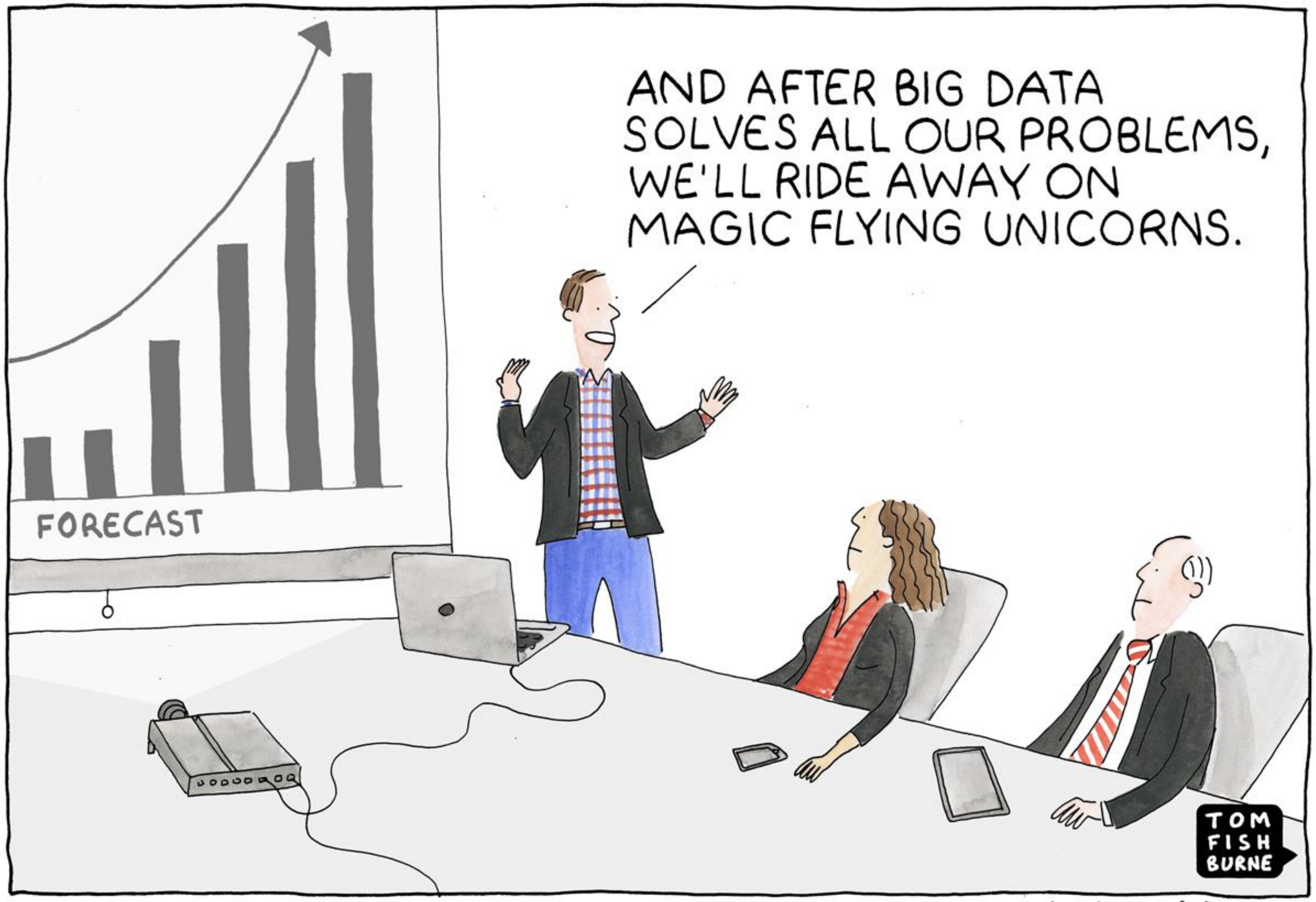
Adjunct Professor, Urology, UCSF

Co-Lead UCSF-KPNC Benign Urology K12

CAIRIBU Meeting
3-4 December 2020
Virtual

Disclosures

- No conflict of interest concerns.



© marketoonist.com

Bigger is (often) Better

- Bigger grants
- Bigger number of outcomes (increased power/precision)
- Bigger number of covariates
- Bigger geospatial area covered
- Bigger number of individuals exposed with wider variation (as well as unexposed)
- Bigger (longer) length of follow-up

What is Big Data Anyway?

- Not large number of rows nor a large number of columns
  - GWAS

- **Big data** is a *field* that treats ways to analyze, systematically extract information from, or otherwise deal with data sets that are *too large or complex* to be dealt with by *traditional* data-processing application software.

- Use of AI/ML in big data
- Will frame this mostly within the context of use of EHR data

What Big Data Is Not or Maybe Traditional is as Better

- "Firstly, routine examination data and definitive diagnosis results for **59 UTI patients gathered and composed as a UTI dataset**. Three classification models namely; decision tree (DT), support vector machine (SVM), random forest (RF) and artificial neural network (ANN), which are widely used in medical diagnosis systems, were created to model the definitive diagnosis results using the composed UTI dataset. Accuracy, specificity and sensitivity statistical measurements were used to determine the performance of created models."   2018 paper.

- A study of 1903 patients looking at prediction of bladder outlet obstruction in men with LUTS.

  "The use of (artificial neural networks), which are better able than traditional regression models to identify non-linear relations and complex interactions between variables, did not improve the prediction of BOO. "

Useful to think about big data on (at least) two dimensions

- Methods
- Applications

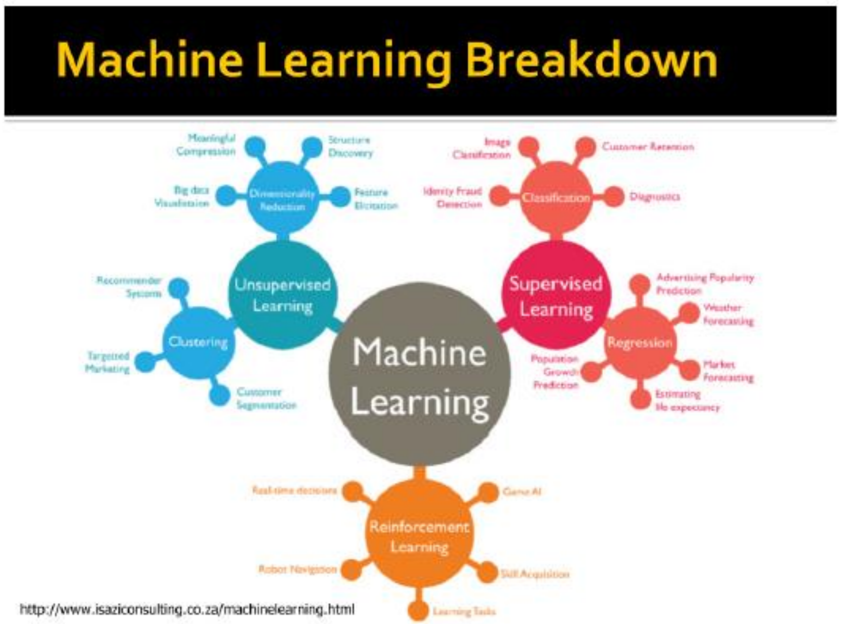


FIGURE 1-3 | A summary of the most common methods and applications for training machine learning algorithms
SOURCE: Reprinted with permission from Isazi Consulting, 2015. http://www.isaziconsulting.co.za/machinelearning.html.

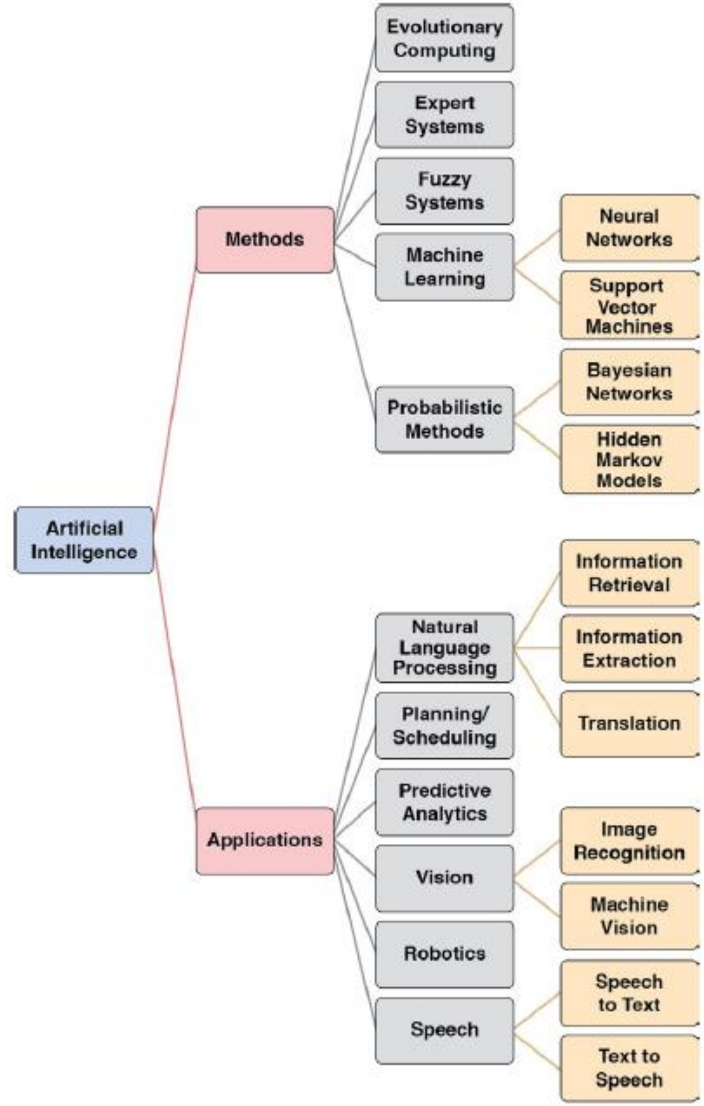


FIGURE 1-2 | A summary of the domains of artificial intelligence
SOURCE: Adapted with permission from a figure in Mills, M. 2015. Artificial Intelligence in Law - The State of Play in 2015? *Legal IT Insider.* https://www.legaltechnology.com/latest-news/artificial-intelligence-in-law-the-state-of-play-in-2015.

Data sources for AI/ML

- EHRs (dx, images, lab tests, provider, locations, etc.)
- Google search
- Omics – genomic, metabolomics,
- Wearables and/or mobile data sourcing attached to an individual
- Patient (or caretaker) reports
- Provider patterns

What is the state of AI/ML in urology

BJUI
BJU International
Reviews

Current status of artificial intelligence applications in urology and their potential to influence clinical practice

Jian Chen[*], Daphne Remulla[*], Jessica H. Nguyen[*], Aastha[†], Yan Liu[†], Prokar Dasgupta[‡] and Andrew J. Hung[*]

Overwhelmingly in urological cancer and radiomics

Recent Examples of ML in Benign Urology

- Performance of a Natural Language Processing Method to Extract Stone Composition From the Electronic Health Record. Bejan CA, et al. Urology. 2019. PMID: 31310771.
- Cardiovascular disease and stroke risk assessment in patients with chronic kidney disease using integration of estimated glomerular filtration rate, ultrasonic image phenotypes, and artificial intelligence: a narrative review. Jamthikar AD, et al. .Int Angiol. 2020 Nov 25. PMID: 33236868
- Development and Validation of a Machine Learning Algorithm for Predicting Response to Anticholinergic Medications for Overactive Bladder Syndrome. Sheyn et al. PMID:31599833
- Supervised Learning Classifiers for Electrical Impedance-based Bladder State Detection. Dunne E, et al. .Sci Rep. 2018 Mar 29;8(1):5363. PMID: 29599451

Additional machine learning efforts

- A framework for diagnosis of urinary incontinence disease based on scoring measures and automatic classifiers. Díaz I, et al. Comput Biol Med. 2011 Jan;41(1):11-7. PMID: 21075362
- Comparison of genetic algorithms and other classification methods in the diagnosis of female urinary incontinence. Laurikkala J, Juhola M, Lammi S, Viikki K. Methods Inf Med. 1999 Jun;38(2):125-31.PMID: 10431517
- Predicting risk of pelvic floor disorders 12 and 20 years after delivery. Jelovsek JE et al. Am J Obstet Gynecol. 2018 Feb;218(2):222.e1-222.e19 PMID: 29056536

Issues to be concerned about

- Data bias
  - Developed in one system & applied in another
    - Differences in patient populations, population stability, provider mix, workflows
  - Absence of groups, types of patients, data
- Data evolution/continuity
  - ICD9 to ICD10
  - New tables/data recorded
  - Data location altered
  - Upgrades
- Validity – replication

Big Data opportunities in benign urology

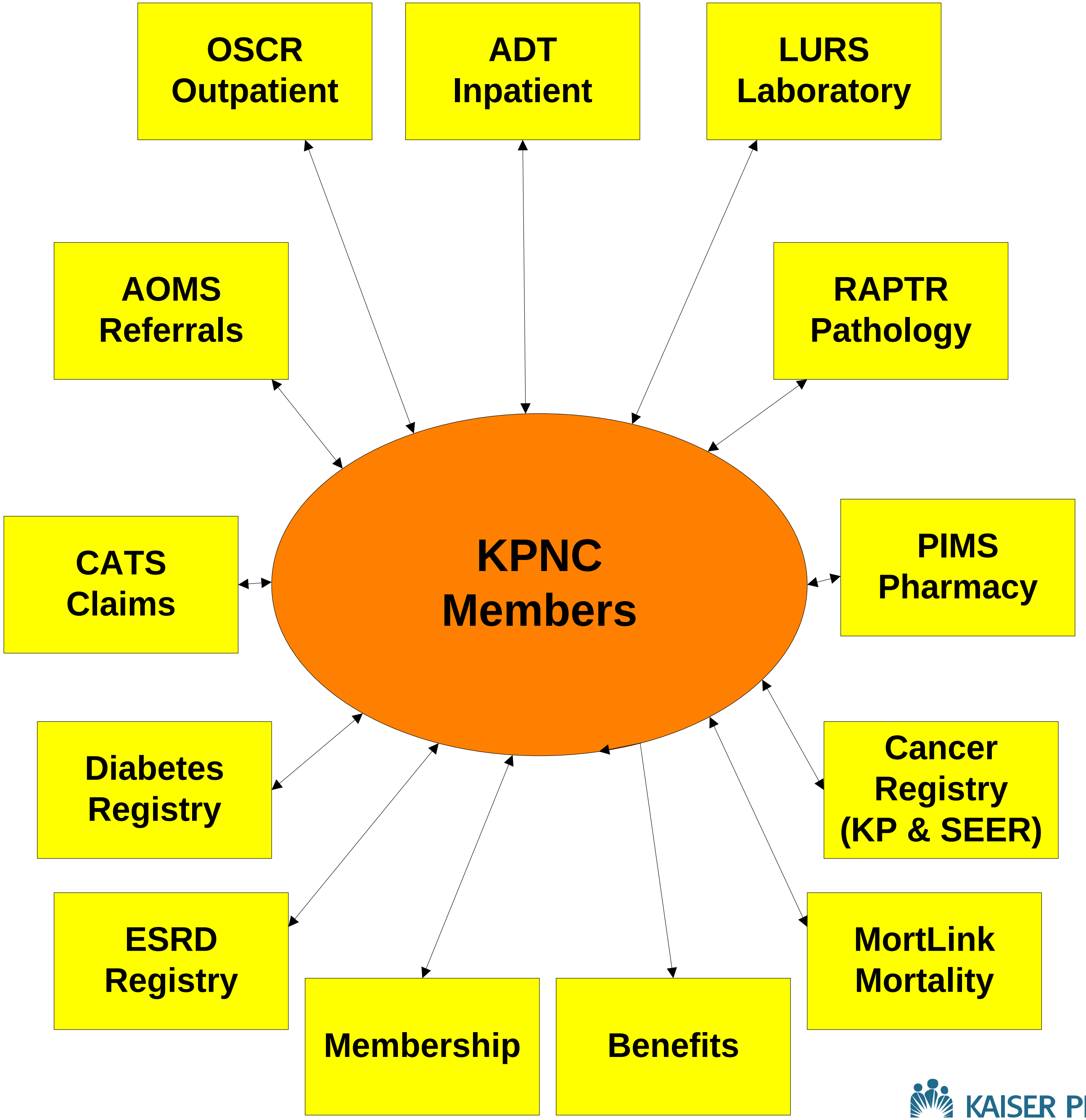
- Predictive analytics
  - Who is high risk? Recurrent stones, progressing to UI,
  - Where can early detection be of benefit?
  - What is the best treatment option for this patient?   Prostatitis, UI, stones
- Natural Language Processing
  - What is buried in text of providers, reports – to identify, add data, screen
- Image processing / radiomics / pathomics
  - What is unseen?   Earlier detection.
- Phenotyping
  - Can we define benign urologic conditions more precisely?

Big Data opportunities in benign urology

- Can we improve our diagnostic process?
  - Better/more careful reading of diagnostic tests
  - Better/more careful assembling and integration of relevant tests & characteristics
  - Better definition of the clinical problem
- Can we improve our treatment decisions?
  - Choosing the right drug or other treatment or the one more likely to yield success
  - Choosing the timing of treatment

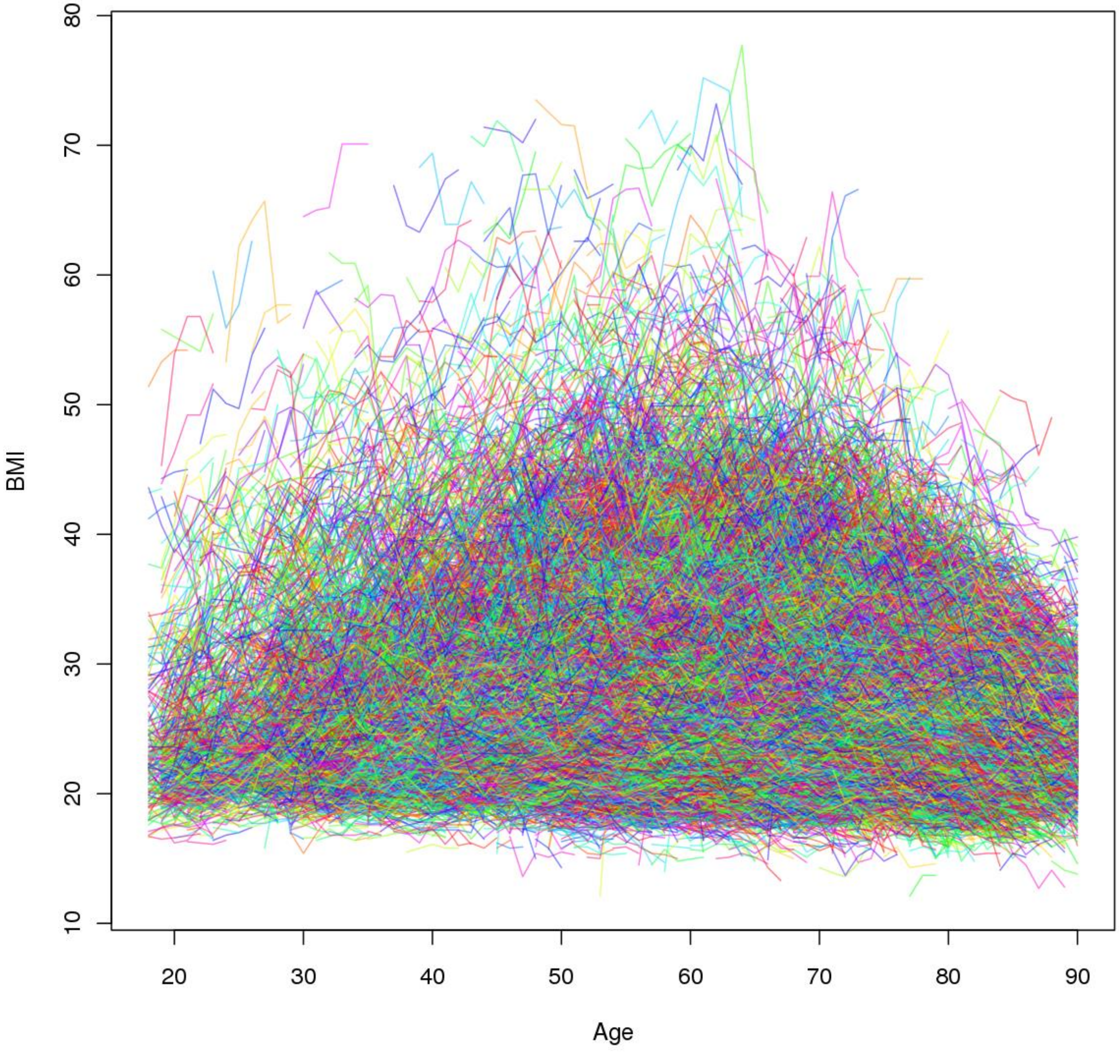
Electronic what?

- EHR, EMR, CPR, AHR…….
  - EHR = electronic health records
  - EMR = electronic medical record
  - CPR = computer-based patient record
  - AHR = automated health records
- For purposes of the talk – any record or set of health records electronically created and stored that replaces paper records
- Comprehensive EHR – all health contacts and information captured and recorded electronically – visits, calls, emails, telemedicine, advice nurse, procedure information, imaging, pathology, pharmacy



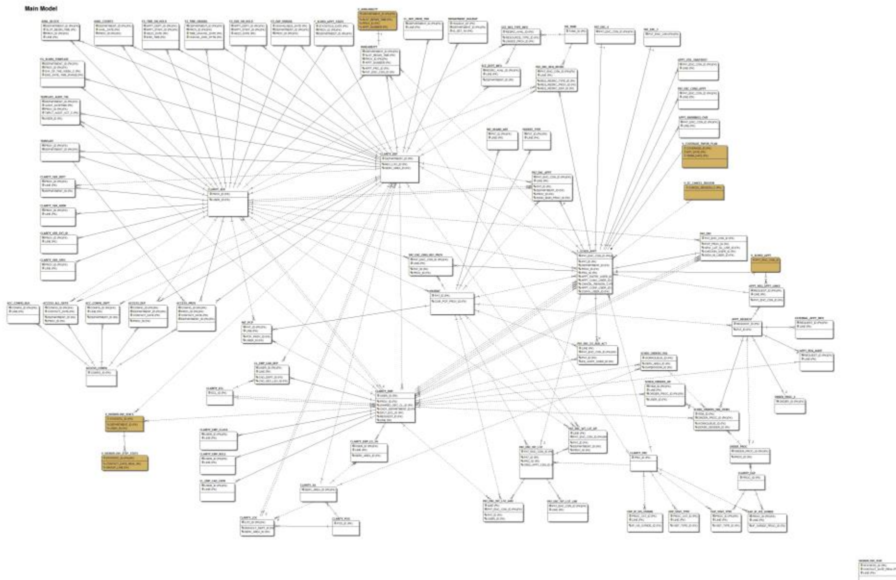
KAISER PERMANENTE®

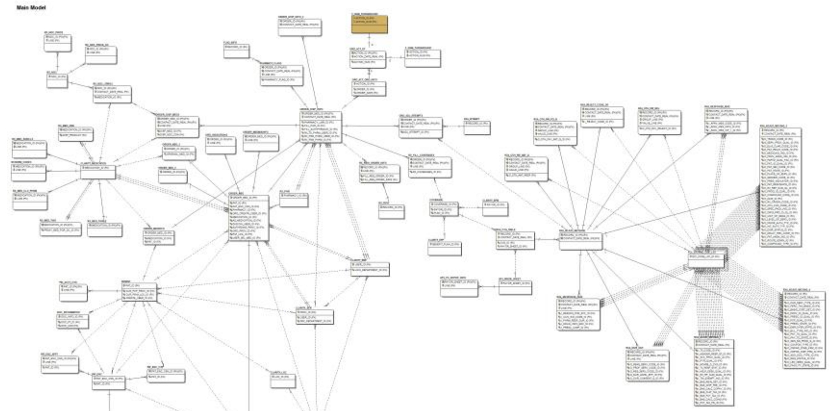
It is easy working with electronic data – you have it all.

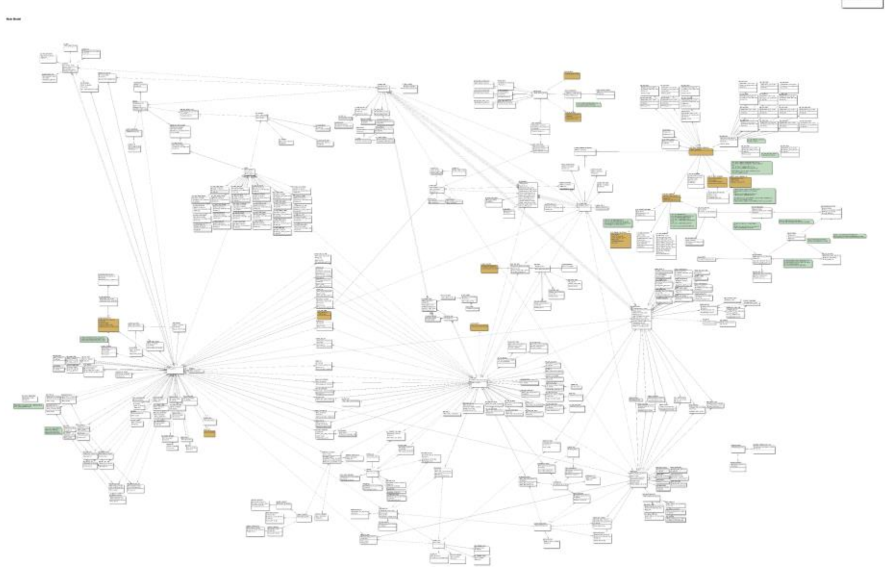


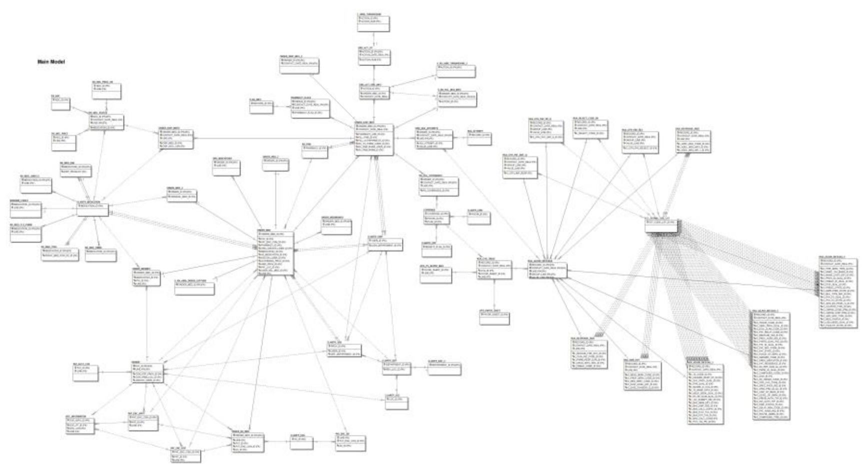
KAISER PERMANENTE®

Complexity

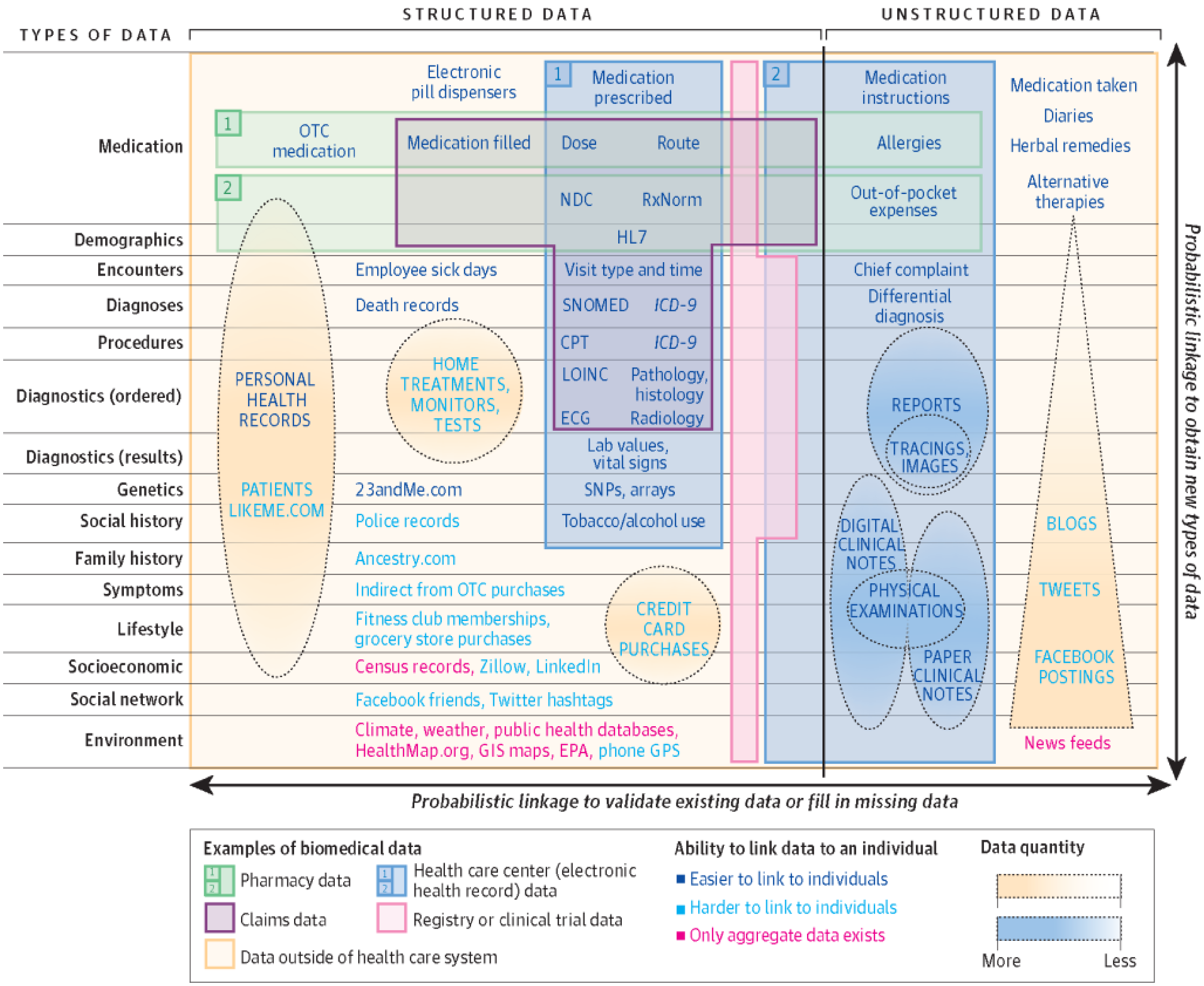








Complexity



NENTE®

Future of EHRs

- Continued rapid growth/penetration
- Continued consolidation
- Significant efforts by computer & health IT scientists (& private companies) to facilitate data merging across systems
- Increasing use of synoptic reporting
- Increased use of free text / natural language processing to derive hard to quantify or qualify health markers
- Increasing capture of non-clinic/hospital health information
  - Telemedicine
  - Device monitoring – glucose, cardiac, neurophysiological
  - Patient symptoms
- Potential fundamental redesigns of EHRs

Our machine learning efforts

- Marvin Langston, PhD
  - Using ML for prediction of patients at high-risk of post-op urinary retention
  - Using ML for prediction of BPH/LUTS progression following treatment (focusing on treatment choice).

CAIRIBU Related Big Data Efforts

CHOP/Penn Center for Machine Learning in Urology (CMLU)

Gregory E. Tasian

Goal is to apply machine learning to improve the understanding of the pathophysiology, diagnosis, risk stratification, and prediction of treatment responses of benign urological disease among children and adults.

Wisconsin Quality-of-Life Machine Learning Algorithm (WISQOL-MLA)

Kris Penniston, Tom Chi

Using ML to predict QoL among individuals with kidney stones. BJU 2020. PMID: 33205549

Other non-random things not mentioned

- Data security
- Repeated evaluations/validations needed
- There can be too much of a good thing
- AI-driven methods for patient assistance (for example, in patients with disabilities)
- Increasing AI involvement in procedures
- Desire to accumulate larger amounts of data vs. a patient's right to decide if their data should be included
- The monetarization of big data/AI
- Level of concern related to false positive (type 1 error) or false negative (type 2 error) varies by what you are trying to do and impact

Thank you.

More information?

Stephen Van Den Eeden, PhD

Stephen.vandeneeden@kp.org

Marvin Langston, PhD

Marvin.E.Langston@kp.org

KAISER PERMANENTE®