The Why and How of Machine Learning for Health Care Systems

Better Predictions, Better Outcomes

Marshall Ruffin, MD,
MPH, MBA, CPE, FAAPL
President and CEO, Salutary, Inc.
Marshall.Ruffin@Salutary.com
The Leadership Institute, Millennium Group, 3/13/2020

Agenda

- Against the Gods, The Remarkable Story of Risk
  - A Brief Summary for Health Care Leaders
- Prediction Machines – Plummeting Costs of Prediction
  - What Machine Learning Permits Us To Do
- The Curse of Dimensionality
  - The Necessary Role of Facilitated Networks of Collaborating Health Care Systems
- Federated Learning and Differential Privacy
  - Protecting Privacy of Individual Patients While Building the Best Predictions
- Changes in Workflow – To Benefit from More Accurate Predictions
  - Equations that Enhance Patient Care and the Value of Clinicians & Managers

Against the Gods, The Remarkable Story of Risk

- Peter Bernstein, John Wiley & Sons, 1996
- “The revolutionary idea that defines the boundary between modern times and the past is the mastery of risk: the notion that the future is more than a whim of the gods and that men and women are not passive before nature. Until human beings discovered a way across that boundary, the future was a mirror of the past or the murky domain of oracles and soothsayers who held a monopoly over knowledge of anticipated events.”
- “The ability to define what may happen in the future and to choose among alternatives lies at the heart of contemporary societies.”

Prediction of Risk

- Relevance for health care organizations
  - Normal distribution and comparison of actual to expected is central to clinical research and identification of special causes.
- “All the tools we use today in risk management and in the analysis of decisions and choice, from the strict rationality of game theory to the challenges of chaos theory, stem from developments that took place between 1654 (Pascal) and 1760 (Bayes).”
  - Thomas Bayes – Bayes Theorem – Spawned Probability Theory – He worked out how to update future probabilities with new information today.
- With dozens of variables, the complexity was unconquerable with arithmetic and calculators. Fast computers & Big Data came to the rescue.

The Curse of Dimensionality

- **Main thesis** – as the number of characteristics grows arithmetically, the number of examples needed grows exponentially, or faster
- Unique distributions of predictors – the dimensions are multiplied:
  - Gender – Male or Female (2 categories) 2
  - Age – 40 – 90 (50 categories) 100
  - Height – 48 to 84 inches (36 categories) 3,600
  - Weight – 50 to 300 pounds (250 categories) 9,000,000
  - Principle Diagnosis – (72,184 ICD10 codes in 2020) 649,656,000,000
  - Genes (normal or not) – (42,000 genes from two parents) 2.7285552e+16
- The number of patients needed to fill each multi-dimensional space quickly becomes astronomical

The Curse of Dimensionality

- To predict health care outcomes: THE MORE PATIENTS THE BETTER!
- No single health care institution has enough patients for best models
  - Not even Kaiser Permanente with 12,500,000 covered lives
- Common mistakes:
  - Create a model on one organization's data and run it on another organization's data. Problem – too few patients.
  - Create a model on merged and de-identified data from many organizations is not optimal, either. Problem – lose valuable detail with de-identified data.

When the Cost of Something Falls, We Use More of It.

- History of the fall in the price of artificial light
  - Candles – expensive, poor light source, people went to bed at dark
  - In 1800 the cost of light was 3,200 times the cost of light today
  - We use more light today than we did in 1800, and light the night
- The rise of the Internet dropped the cost of distribution, communication and search – dramatically!
  - Google Searches today: 63,000/second, 3.8M/minute, 228M/hour
- The cost of prediction is falling very fast – so we will use more of it in all industries and all endeavors.

The Cost of Prediction Is Plummeting

- Why is Cost of Prediction Plummeting?
  - Falling costs of hardware (GPUs and memory and processors)
  - Rapidly growing data from electronic medical records & genetics
  - New generation of data management tools (Hadoop, Spark, etc.)
  - Deep learning software (Python, TensorFlow, etc.)
- What societal changes occur when cost of prediction falls?
  - We automate more prediction
  - We elevate the importance of people whose jobs are supported by prediction – we elevate human judgment.

From AI Winter to AI Summer – Why? How?

- Why Did AI Winters from 1970s through 1990s Occur?
  - AI with orderly rational rules ("Expert Systems" using symbolic expression of rules, or IF-THEN statements, in LISP) and was effective with only the simplest problems and did not scale to complex problems
- How Did the AI Summer of 2010's Occur?
  - AI adopted pattern recognition with fast computers and many examples – patterns in data with Python & Hadoop

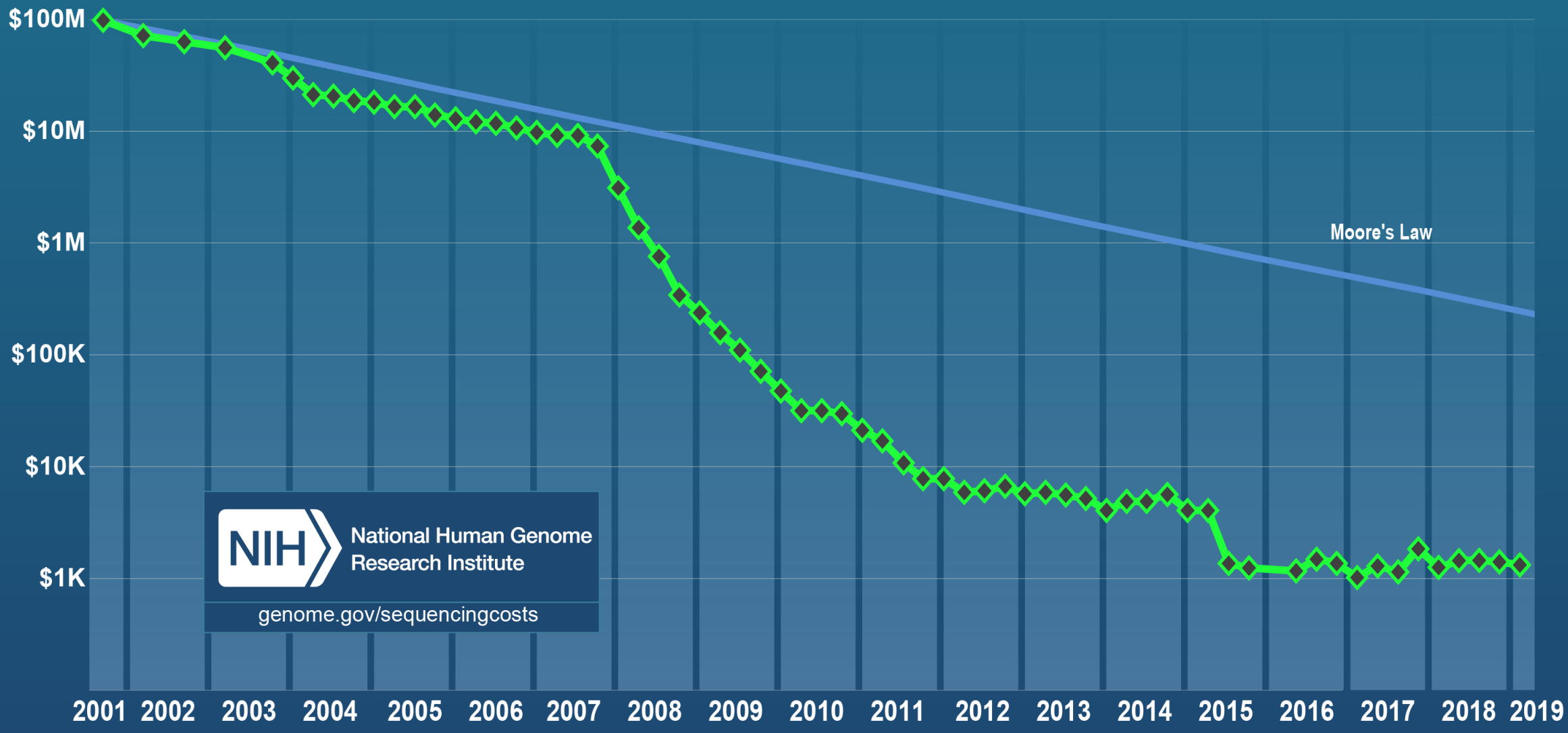
Prediction Machines, The Simple Economics of Artificial Intelligence

- Agrawal, Gans and Goldfarb, Harvard Business Review Press, 2018.
- “Artificial intelligence (AI) is a prediction technology, predictions are inputs to decision making, and economics provides a perfect framework for understanding the trade-offs underlying any decision.”
- “What will new AI technologies make so inexpensive? Prediction. Therefore, as economics tells us, not only are we going to start using a lot more prediction, but we are going to see it emerge in surprising new places.”

Why is Accurate Prediction Important?

- Better prediction reduces uncertainty & permits more efficient & effective use of resources to secure desirable outcomes.
- In other words, better predictions produce better outcomes.
- Better prediction in health care – more accurate diagnoses.
- Better prediction in automobiles – safer self-driving.
- Better prediction in airlines – safer flights.
- Better prediction in electrical grid – fewer blackouts, lower costs.
- Humans plus algorithms are more accurate than either alone.

Cost per Genome



Now Effective Prediction is All Around Us

- Whole Exome and Genome Sequencing,
- Google and Bing search engines,
- Effective language translation,
- Voice and face recognition – Google Translate & IBM Watson Face Recog.,
- Automated assistants such as Siri and Alexa,
- Optimizing electricity grids,
- Automated drug discovery,
- Airplane auto-pilots,
- Self-driving trucks (mining) and carts (warehouses) and cars,
- Computers that play checkers, chess, and Go better than humans.

32 Examples of AI in Health Care

- "32 Examples of AI in Health Care," Sam Daley, https://builtin.com/artificial-intelligence/artificial-intelligence-healthcare, www.builtin.com, February 24, 2020.
- AI and Image Interpretation
  - Radiology & Pathology: PathAI, Buoy Health, Enlitic, Zebra Medical Vision, etc.
- Drug Development
  - Bioxcel Therapeutics, Berg Health, Xtalpi, Atomwise, Deep Genomics, etc.
- Improving Patient Experience
  - Olive, Qventus, Babylon Health, CLOUDMEDX, IBM at Cleveland Clinic, etc.
- Mining and Managing Medical Data
  - Tempus, KENSCI, PROSCIA, H20.ai, IBM Watson, Google Deepmind Health, etc.
- AI Robot-Assisted Surgery
  - Vicarious Surgical, Auris Health, Accuray, Intuitive, Heartlander, Microsure, Mazor Robotics

Examples of Predictions Important in Health Care and Value of Genetics to Them

- 2007 – 240 papers on genome wide association studies (GWAS) included < 500 people because genome sequencing was so expensive.
- 2019 – 3,700 papers on GWAS included hundreds of thousands to millions of people because of the falling price of genome sequencing
- Early results
  - 2007: genes associated with age-related macular degeneration AUROC 82%
- Recent results
  - 2018: genes associated with Alzheimer's disease AUROC 84%

How Does One Predict Outcomes and Adhere to HIPAA Guidelines for Privacy?

- HIPAA Primary Use
- Federated Learning
- Differential Privacy

Federated Learning

- Train deep learning models across separate databases without merging their data.
- Move weights of predictors and not data themselves.
- Do not de-identify the data, better for primary use.
- Preserve dates & times of service, location and identities of providers for more accurate models.

Differential Privacy

- Wikipedia: Differential Privacy is a system for publicly sharing information about a dataset by describing the patterns of groups within a dataset while withholding information about individuals in the data set.
  - Adding “mathematical noise” into the data
- Used by Google, Amazon and Apple to protect privacy
- Adopted by US Government for 2020 Census

How to Change Workflow for Predictions?

- More accurate predictions lead to more prudent judgments
- More prudent judgments lead to more successful decisions
- Push predictions into workflow of those who make decisions when they need to make those decisions
  - Who makes specific decision(s)? Physician? Nurse? Manager?
  - What decision(s) does that person make?
  - When does that person make the decision(s)?
  - Where does that person make those specific decision(s)?
- Workflow may need redesign to use new predictions best!

Summary

- Risk management needs accurate predictions of future events.
- Prediction of risk in health care ia challenged by Curse of Dimensionality.
- The more data (+ genetics) and the more patients' records the better.
- The price of prediction is falling.
- We will use more prediction in the future.
- Prediction is more accurate with data from multiple organizations.
- Federated learning and differential privacy are essential.
- Workflow may change to benefit from more accurate predictions.