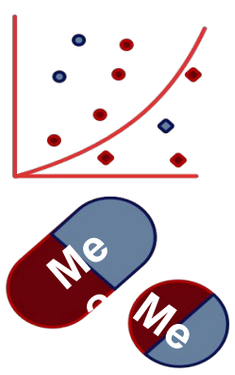
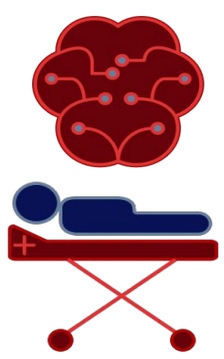
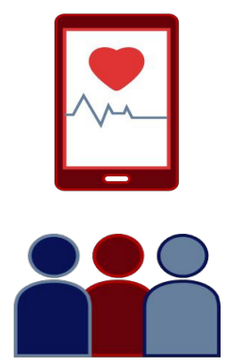
# Learning Healthy Models for Healthcare

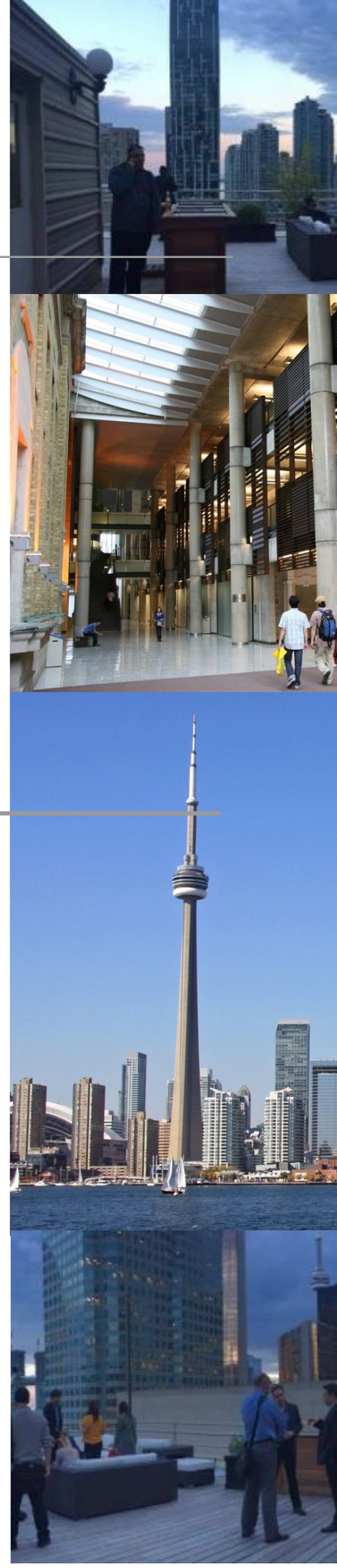
Dr. Marzyeh Ghassemi, PhD MIT CSAIL
University of Toronto, Computer Science and Medicine
Canadian CIFAR AI Chair, Vector Institute
NeuroIPS Workshop Co-Chair

www.marzyehghassemi.com, @MarzyehGhassemi



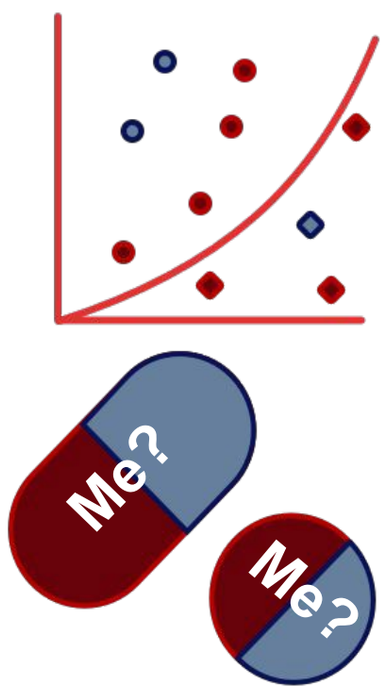





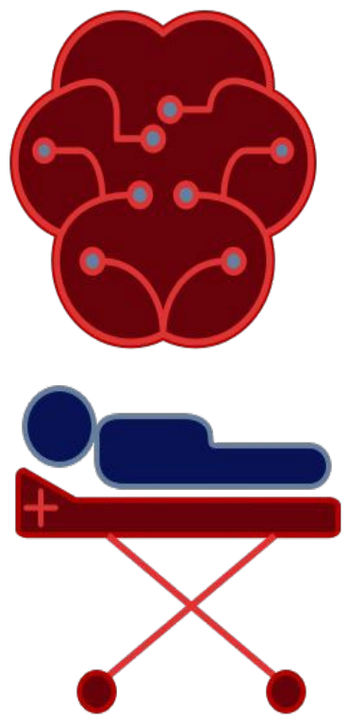


# Machine Learning For Health (ML4H)

**Creating actionable insights in human health.**



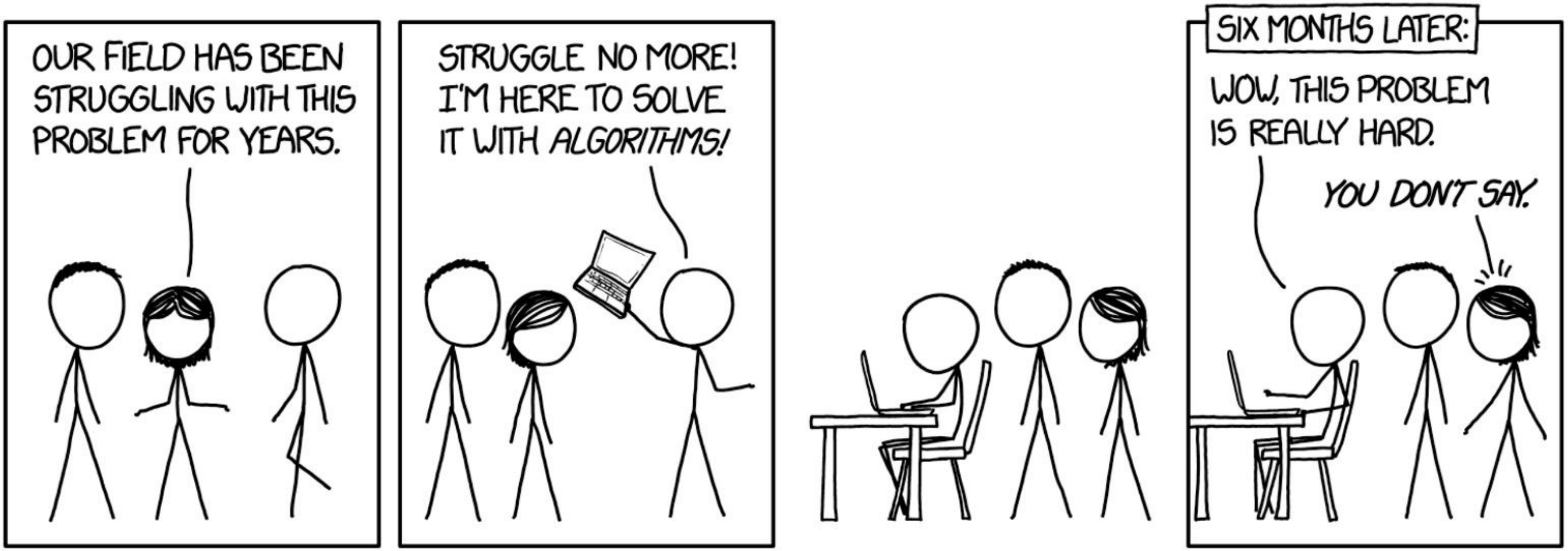

What models are healthy?



What healthcare is healthy?

What behaviors are healthy?

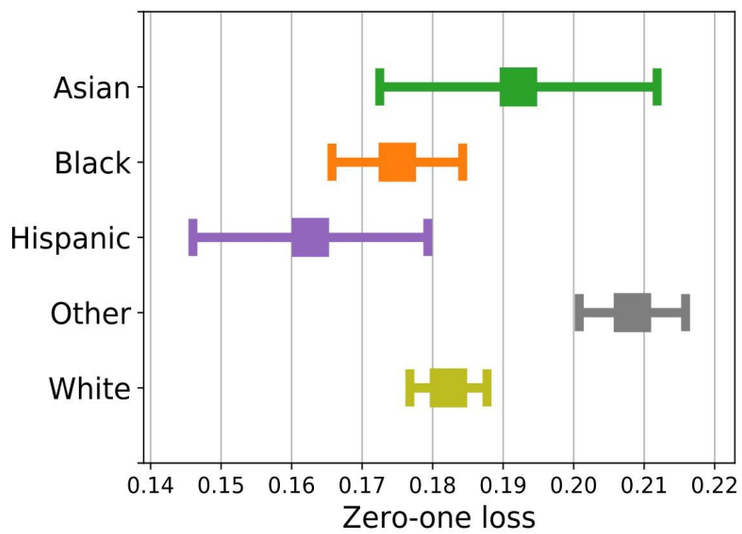
# Why Do We Need Transparency In Models?

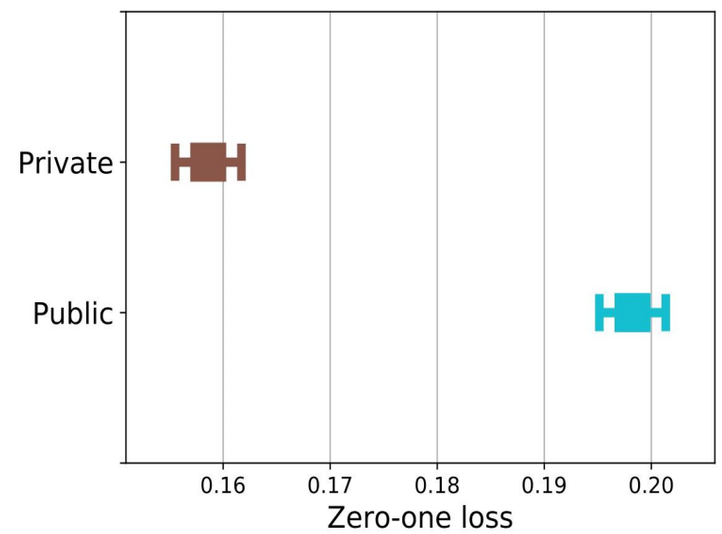


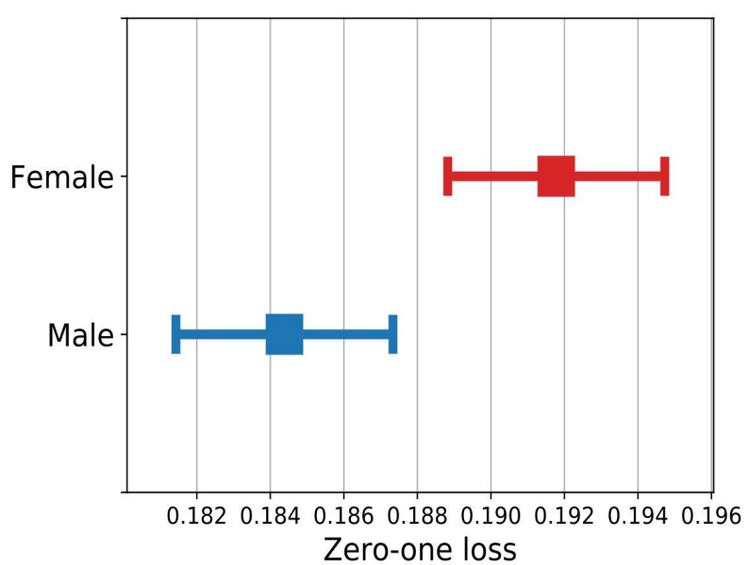

Quelle: XKCD by Randall Munroe

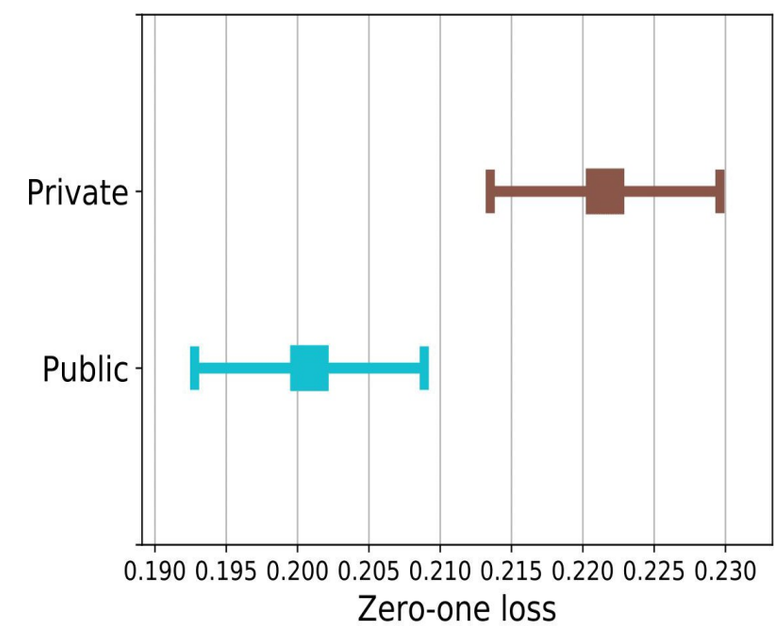
# **Fairness** in Medical and Mental Health Accuracies

- Significant differences in model accuracy for race, sex, and insurance type in **ICU notes** and insurance type in **psychiatric notes**.




[1] Chen, Irene Y., Peter Szolovits, and Marzyeh Ghassemi. "Can AI Help Reduce Disparities in General Medical and Mental Health Care?." *AMA journal of ethics* 21.2 (2019 167-179.

# Ethics in Health Is Not New; But **Ethical ML** Is.

- **Clinical trial populations:** Clinical Trials Still Don't Reflect the Diversity of America (NPR, Dec 2015)

- **Retracted studies:** Harvard Calls for Retraction of Dozens of Studies by Noted Cardiac Researcher (NYT, Oct 2018)

- **Conflict of interest:** Sloan Kettering's Cozy Deal with Start-Up Ignites a New Uproar (NYT, Sept 2018)

# Even With Ethics Training, Bias Is Part of the Clinical Landscape

- How does/should ML interact with fairness/health[1,2,3,4,5]?

This Issue | Views 12,435 | Citations 41 | Altmetric 174

Viewpoint

August 11, 2015

**Racial Bias in Health Care and Health Challenges and Opportunities**

David R. Williams, PhD, MPH[1,2]; Ronald Wyatt, MD, MHA[3]

Author Affiliations

*JAMA*. 2015;314(6):555-556. doi:10.1001/jama.2015.9260

J Palliat Med. 2013 Nov; 16(11): 1329–1334.
doi: 10.1089/jpm.2013.9468

PMCID: PMC3822363
PMID: 24073685

Racial and Ethnic Disparities in Palliative Care

Kimberly S. Johnson, MD, MHS[1,2]

Author information ▶ Article notes ▶ Copyright and License information ▶ Disclaimer

This article has been cited by other articles in PMC.

2

The Girl Who Cried Pain: A Bias Against Women in the Treatment of Pain

Diane E. Hoffmann and Anita J. Tarzian

Am J Public Health. 2007 February; 97(2): 247–251.
doi: 10.2105/AJPH.2005.072975

PMCID: PMC1781382
PMID: 17194867

The Black–White Disparity in Pregnancy-Related Mortality From 5 Conditions: Differences in Prevalence and Case-Fatality Rates

Myra J. Tucker, BSN, MPH, Cynthia J. Berg, MD, MPH, William M. Callaghan, MD, MPH, and Jason Hsia, PhD

Author information ▶ Article notes ▶ Copyright and License information ▶ Disclaimer

Obes Rev. 2015 Apr;16(4):319-26. doi: 10.1111/obr.12266. Epub 2015 Mar 5.

**Impact of weight bias and stigma on quality of care and outcomes for patients with obesity.**

Phelan SM[1], Burgess DJ, Yeazel MW, Hellerstedt WL, Griffin JM, van Ryn M.

Author information

1 Continuous State-Space Models for Optimal Sepsis Treatment - Deep Reinforcement Learning … (MLHC/JMLR 2017);
2 Modeling Mistrust in End-of-Life Care (MLHC 2018/FATML 2018 Workshop);
3 The Disparate Impacts of Medical and Mental Health with AI. (AMA Journal of Ethics 2019);
4 ClinicalVis Project with Google Brain. (*In submission);



UNIVERSITY OF TORONTO

5

VECTOR INSTITUTE | INSTITUT VECTEUR

# Human Transparency?

- Human **decisions** about routine practice are **justified** by/in **research**.

**But**

- ~400 **routine practices** were **contradicted** by studies published in leading journals[1].

- More than 10% of 3,000+ studies in JAMA/Lancet/NEJM were a "**medical reversal**": a conclusion opposite of what had been conventional wisdom among doctors[1].

[1] Meta-Research: A comprehensive review of randomized clinical trials in three medical journals reveals 396 medical reversals. Herrera-Perez, Diana, et al. eLife 8 (2019): e45183 [https://elifesciences.org/articles/45183]

# What Kind Of **Transparency** Do We **Need**?

- Important when clinical opinion **differs** significantly from predictions.

- **Doesn't** mean:
  - Clinicians care about how a model works
  - You build simple models for the sake of transparency

- **Could** mean:
  - Understanding when to rely or reject model output
  - Notification of which populations the model might work/fail on
  - Including limitations of data models are trained

What are technical options?

# **Option 1)** Transparency Via Post-hoc Explanations

- Post-hoc explanations are useful if they are **consistent**.
  - different model behavior -> different explanation
  - same model behavior-> same explanation

- Attention based explanations[1] and saliency maps[2] are not always consistent.

- Possible fixes are **competition-based**; e.g., computing maps for all possible labels and use simple competition to remove less relevant pixels[3].

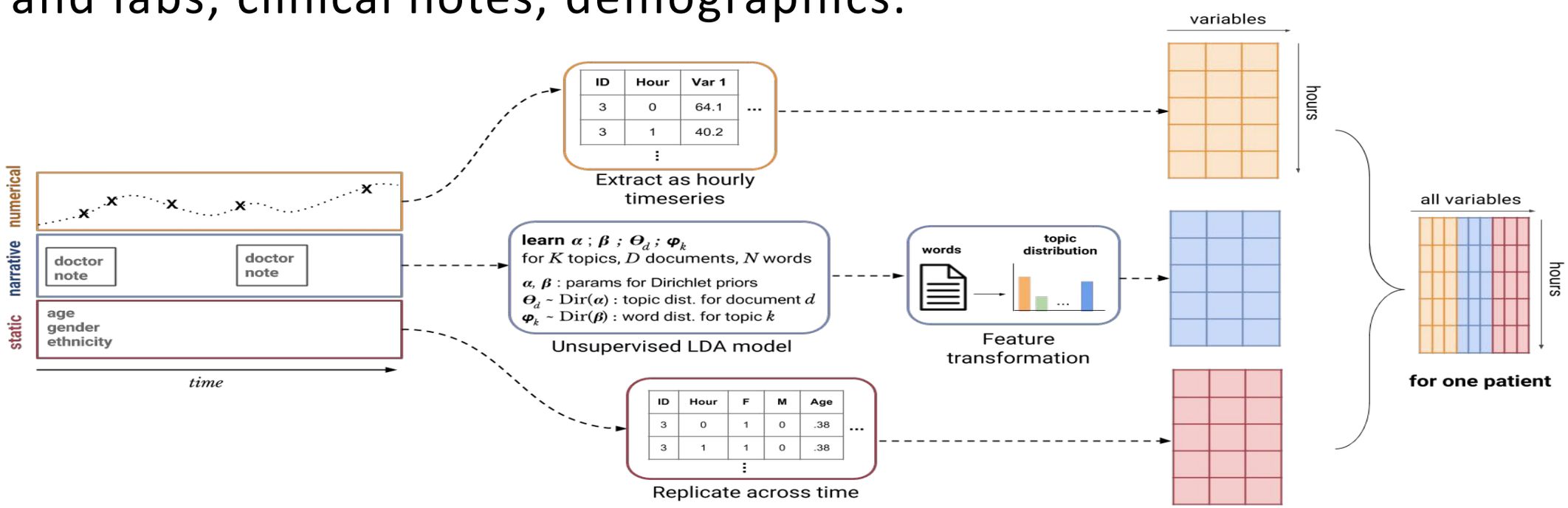
1 Attention is not Explanation. Sarthak Jain, Byron C. Wallace. [https://arxiv.org/abs/1902.10186]
2Sanity checks for saliency maps. Adebayo, J., Gilmer, J., Muelly, M., Goodfellow, I., Hardt, M., & Kim, B. (2018). In Advances in Neural Information Processing Systems (pp. 9505-9515).
3 A Simple Saliency Method That Passes the Sanity Checks. Gupta, A., & Arora, S. (2019). [https://arxiv.org/abs/1905.12152]


8
UNIVERSITY OF TORONTO
VECTOR INSTITUTE | INSTITUT VECTEUR

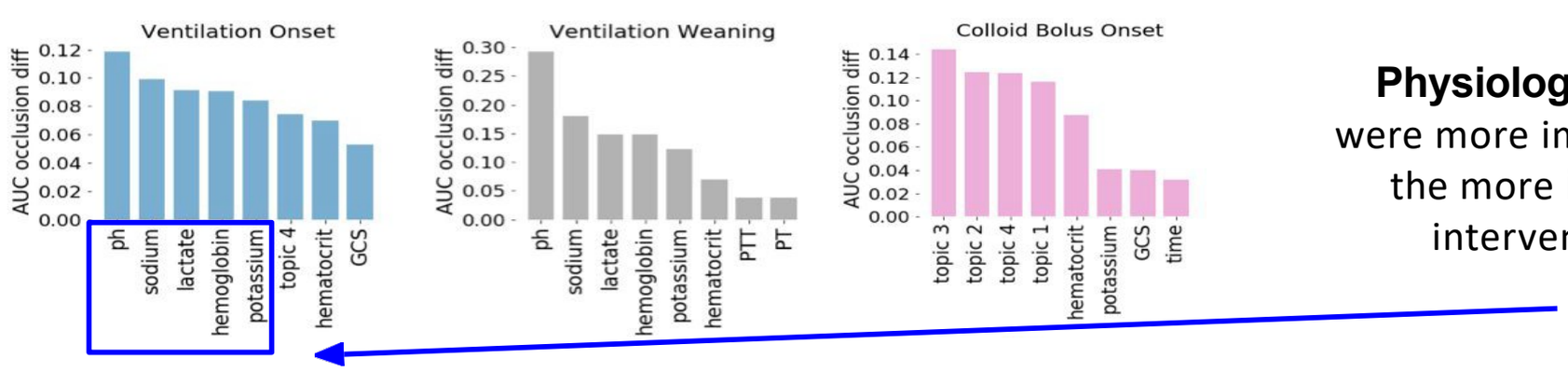
# Clinical Intervention Prediction and Understanding Using Deep Networks

Harini Suresh, Nathan Hunt, Alistair Johnson, Leo Anthony Celi, Peter Szolovits, Marzyeh Ghassemi.
In Proceedings of Machine Learning for Healthcare 2017, JMLR WC Track V68

- Predicting **interventions** for 34,148 ICU patients' time-varying vitals and labs, clinical notes, demographics.




- Feature-level **occlusions** to identify **importance** of information

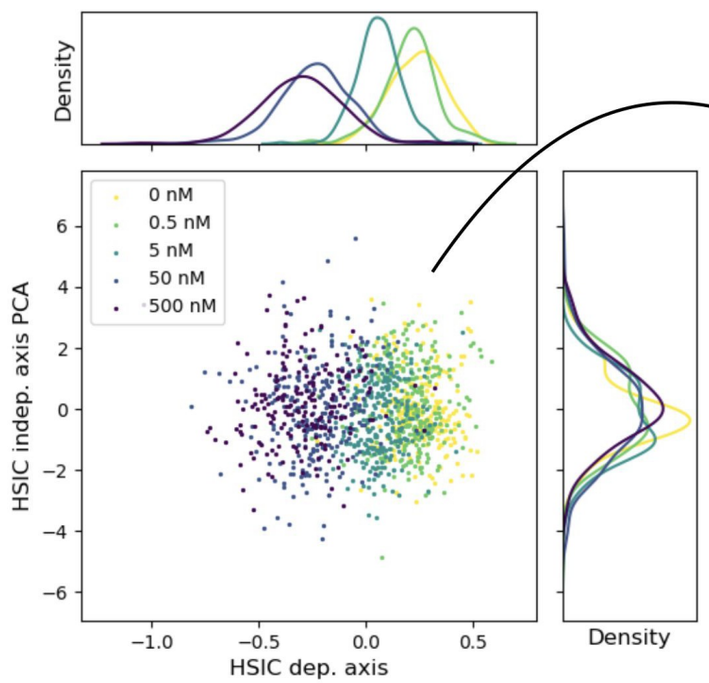


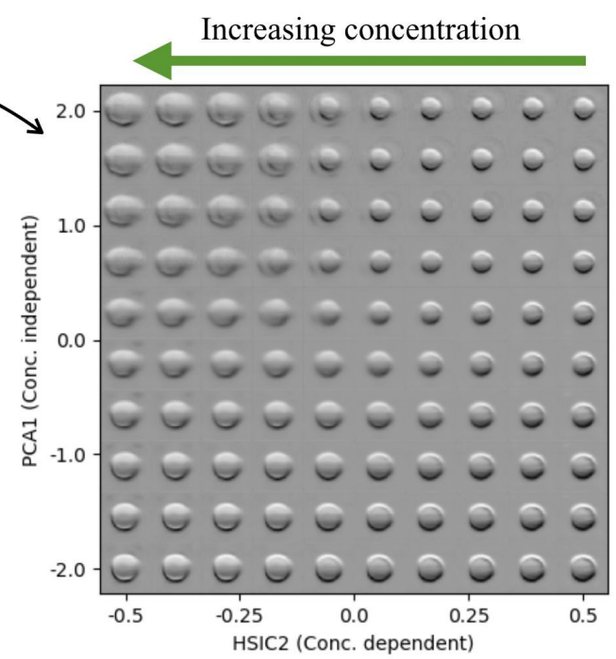

**Physiological data** were more important for the more **invasive** interventions.

# Modeling the Biological Pathology Continuum with HSIC-regularized Wasserstein Auto-encoders

Denny Wu, Hirofumi Kobayashi, Charles Ding, Lei Cheng, Keisuke Goda, Marzyeh Ghassemi
In NeuroIPS 2018 Machine Learning for Health (ML4H) Workshop;

- Regularized generative model for “transparent” latent features; create latent representations that model pathology continuum.




Plot test images on latent space of ~10,000 images from leukemia cell line K562 with dilutions of adriamycin.

Test images show class separation on **x** (dependant axis), but not on **y** (1st PC of independent axes).

Generated images sampled from the dependent axis and the 1st PC of all other axes; generated cells vary in shape.

- HSIC enforces dependency so that latent dimension models continuous **morphological change** corresponding to provided **side information**.

# **Option 2)** Transparency Through "Natural" Outputs

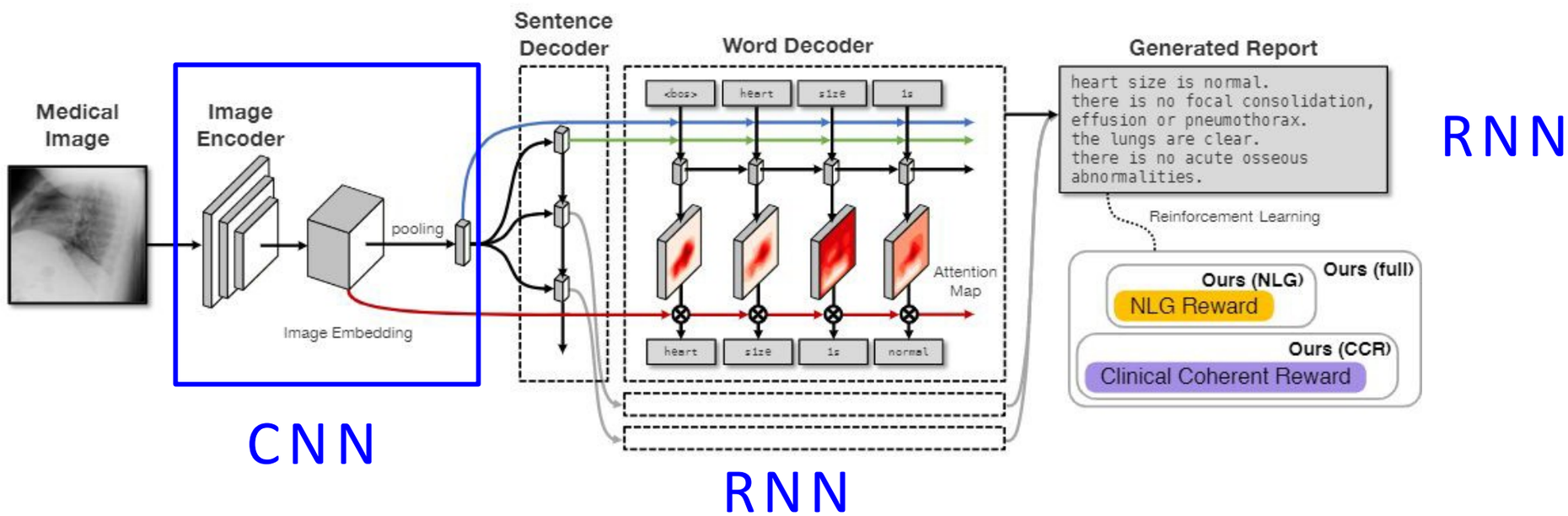
- Transparency is just a way of calibrating **clinicians' trust** in the model.
- Have clinicians and ML experts agree on a set of metrics that are a consumable end product, not an intermediate.
- Delivering these metrics gives transparency.

# Clinically Accurate Chest X-Ray Report Generation

Guanxiong Liu, Tzu-Ming Harry Hsu, Matthew McDermott, Willie Boag, Wei-Hung Weng, Peter Szolovits, Marzyeh Ghassemi.
In Proceedings of Machine Learning for Healthcare 2019, JMLR WC Track

- Automatically **generate** radiology **reports** given medical **radiographs**.

- Chest X-Ray radiology report generation:
  - First predict the **topics** discussed in the report.
  - **Conditionally** generate **sentences** corresponding to these topics.




- CNN-RNN-RNN structure gives model the ability to **use largely templated sentences** and **generate diverse text**.

# **Option 3)** Transparency In **Audit** of Em**bodied** Data

- All data is valuable; embodied health data particularly so.

- Transparent algorithms require large scale datasets for research use.

AWS Machine Learning Blog

**Improving Patient Care with Machine Learning At Beth Israel Deaconess Medical Center**

by Dr. Matt Wood | on 04 MAR 2019 | Permalink | Comments | Share

Beth Israel Deaconess Medical Center has launched a multi-year, innovative research program on how machine learning can improve patient care, supported by an academic research sponsorship grant from AWS. The Harvard Medical School-affiliated teaching hospital will use a broad array of AWS machine learning services to uncover new ways that machine learning technology can enhance clinical care, streamline operations, and eliminate waste, with the goal of improving patient care and quality of life.

**Improving patient care with machine learning**

Inefficiencies in hospital management and operations are not only extremely costly to providers, insurers, patients, and taxpayers, but they can result in precious resources being diverted away from patient care. These inefficiencies drive healthcare costs up and can contribute to life-threatening medical

**Amazon Comprehend Medical**

Extract information from unstructured medical text accurately and quickly

No machine learning experience required

Get started with Amazon Comprehend Medical

Amazon Comprehend Medical is a natural language processing service that makes it easy to use machine learning to extract relevant medical information from unstructured text. Using Amazon Comprehend Medical, you can quickly and accurately gather information, such as medical condition, medication, dosage, strength, and frequency from a variety of sources like doctors' notes, clinical trial reports, and patient health records.

**Google Tries to Patent Healthcare Deep Learning, EHR Analytics**

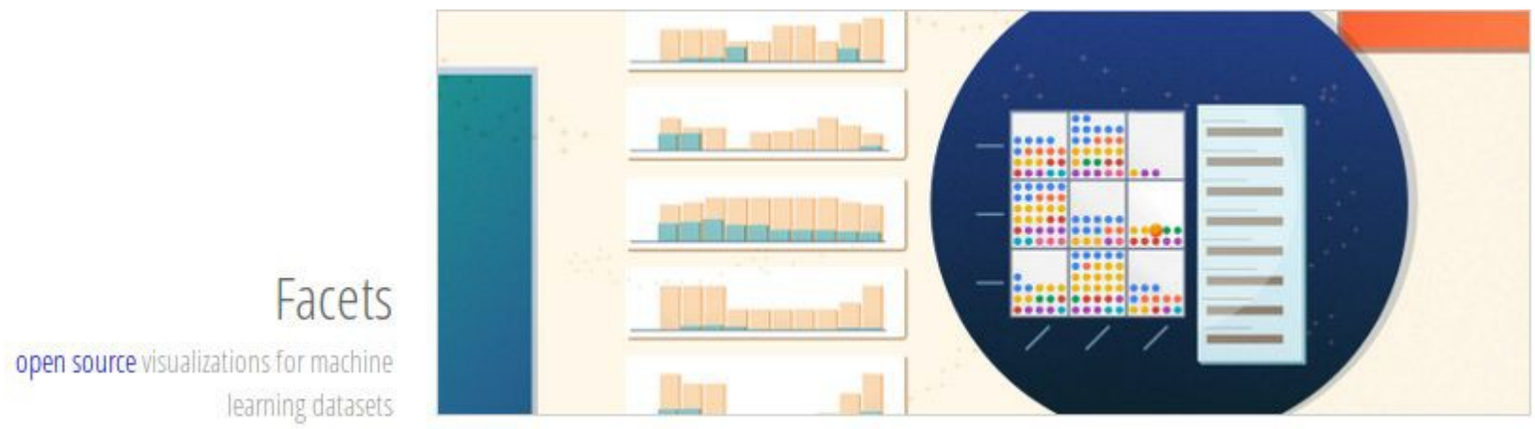
Google has applied for a sweeping patent including the fundamentals of deep learning and EHR analytics in the healthcare industry.

Source: Google

# New Tools to Understand the Data and Process

- Tools like Datasheets[1] for datasets and Modelcards[2] for **model reporting** can be used in a data/model/training agnostic way.

- "Big Picture" tools to **understand potential** data biases.



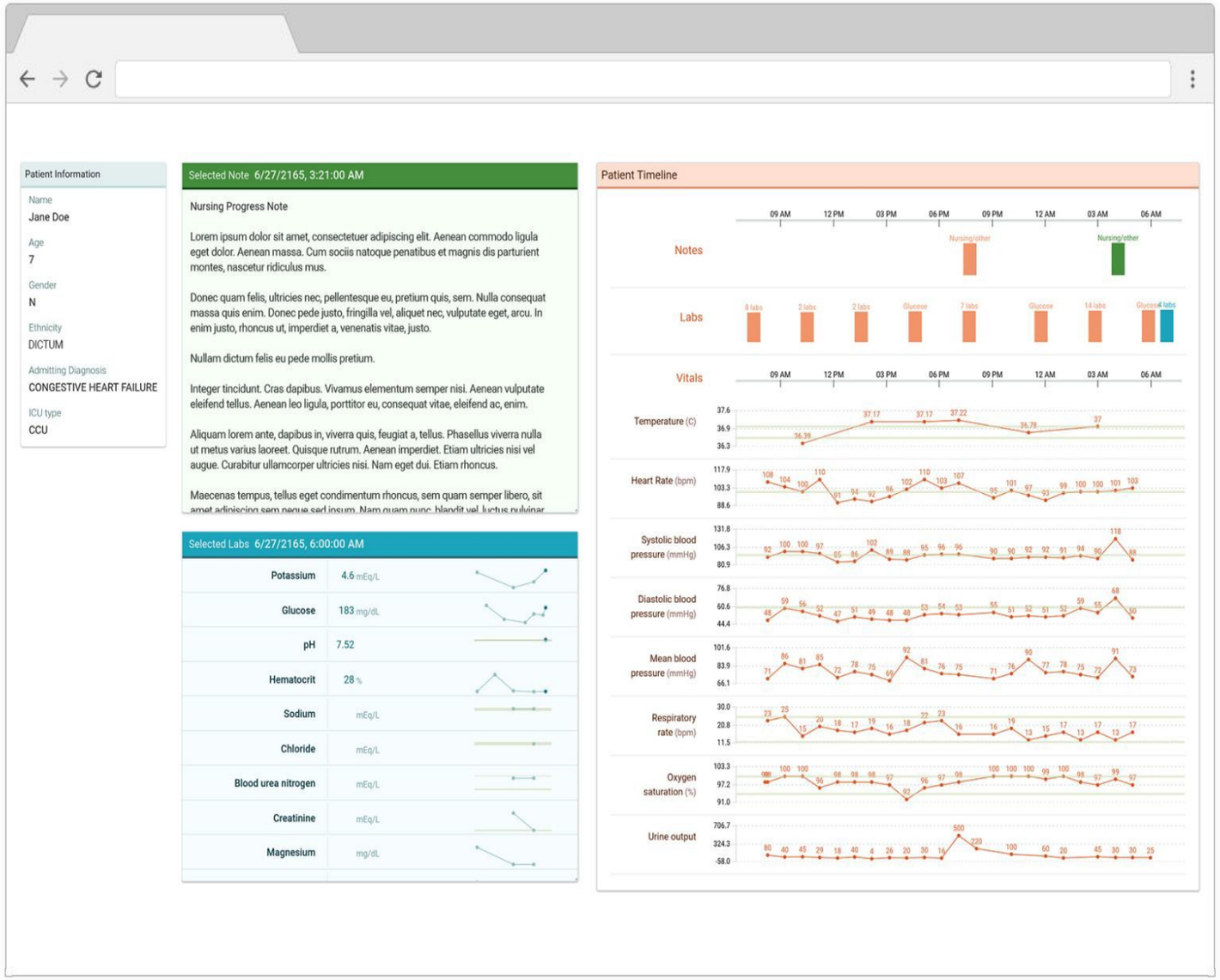

- Working towards **transparent processes**, not models.

1 Datasheets for datasets. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumeé III, H., & Crawford, K. (2018). arXiv preprint arXiv:1803.09010.
2Model cards for model reporting. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019, January). In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220-229). ACM.
[3] https://research.google.com/bigpicture/attacking-discrimination-in-ml/

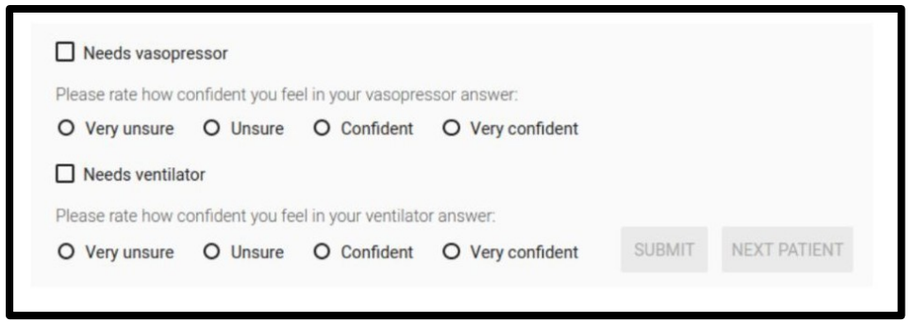
# ClinicalVis: Supporting Clinical Task-Focused Design Evaluation

Marzyeh Ghassemi, Mahima Pushkarna, James Wexler, Jesse Johnson, Paul Varghese

## 1. Present real patient data to HCPs using open-source prototype.




## 2. Ask HCPs to plan care for two interventions in an eICU simulation.




## 3. Evaluate the confidence, accuracy and time-to-task under different visual prototypes.

| | | Vasopressor Positive (VP+) | Ventilator Positive (VE+) |
|---|---|---|---|
| Accuracy (%) | Baseline | 50.00 % | 56.25 % |
| | ClinicalVis | 68.83 % | 62.79 % |
| Confidence Score | Baseline | 0.68 | 0.87 |
| | ClinicalVis | 1.41 | 1.27 |
| Average Time to Task (seconds) | Baseline | 92.31 s | 92.73 s |
| | ClinicalVis | 84.43 s | 86.86 s |

Demo Live at: https://pair-code.github.io/clinical-vis/

UNIVERSITY OF TORONTO

VECTOR INSTITUTE | INSTITUT VECTEUR



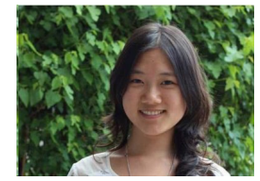
15

# ML4H University of Toronto The Vector Institute

PhD Students

Bret Nestor

Denny Wu


Amy Lu

Matthew McDermott

Technical Collaborators

Dr. Anna Goldenburg

Dr. Shalmali Joshi

Clinical Collaborators

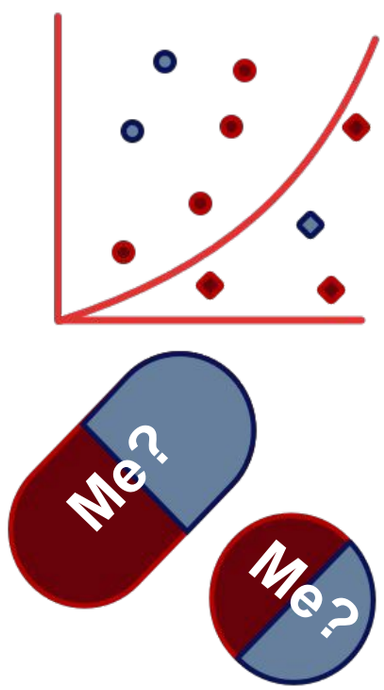
Dr. Amol Verma

Dr. Fahad Razak

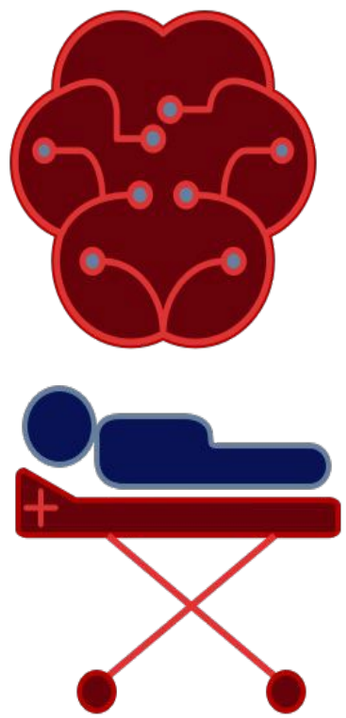
Dr. Muhammad Mamdani

# Machine Learning For Health (ML4H)

**Creating actionable insights in human health.**




What models are healthy?



What healthcare is healthy?

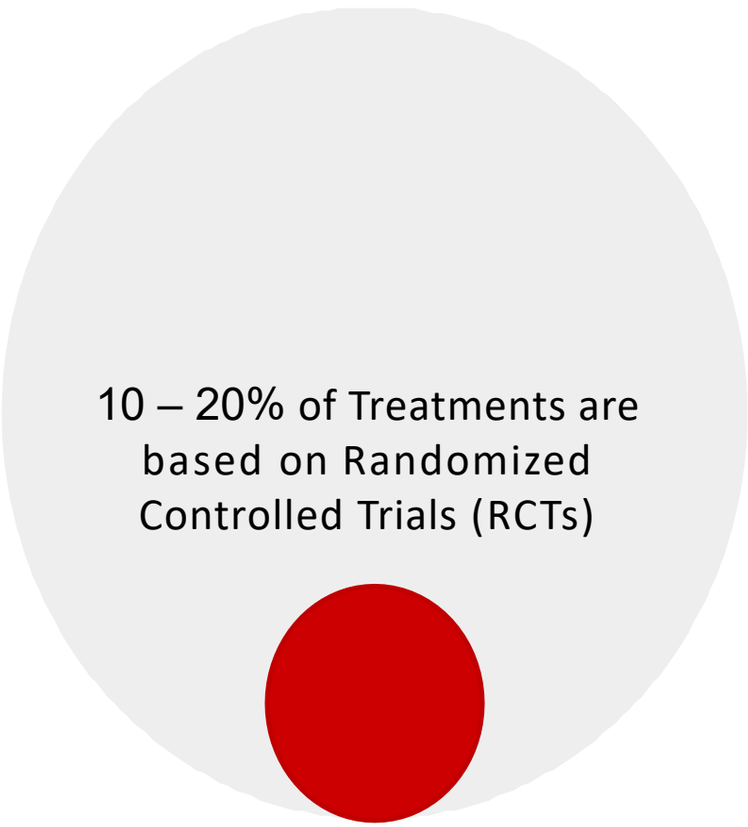
What behaviors are healthy?

# Evidence in Healthcare and Health?

Randomized Controlled Trials (RCTs) are

# Evidence in Healthcare and Health?

Randomized Controlled Trials (RCTs) are **rare and expensive**




[1] *Smith M, Saunders R, Stuckhardt L, McGinnis JM, Committee on the Learning Health Care System in America, Institute of Medicine. Best Care At Lower Cost: The Path To Continuously Learning Health Care In America. Washington: National Academies Press; 2013..*

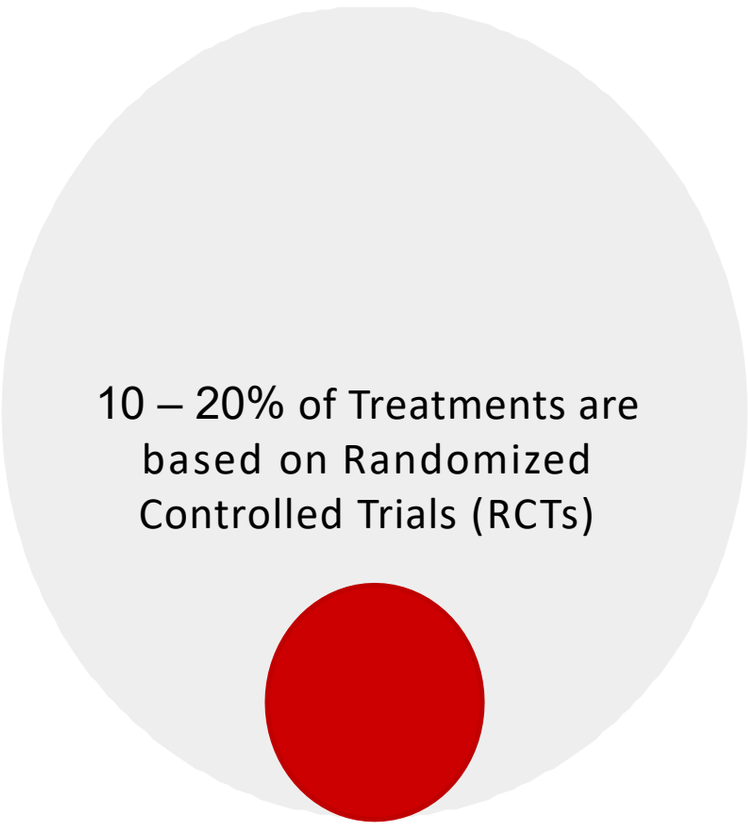


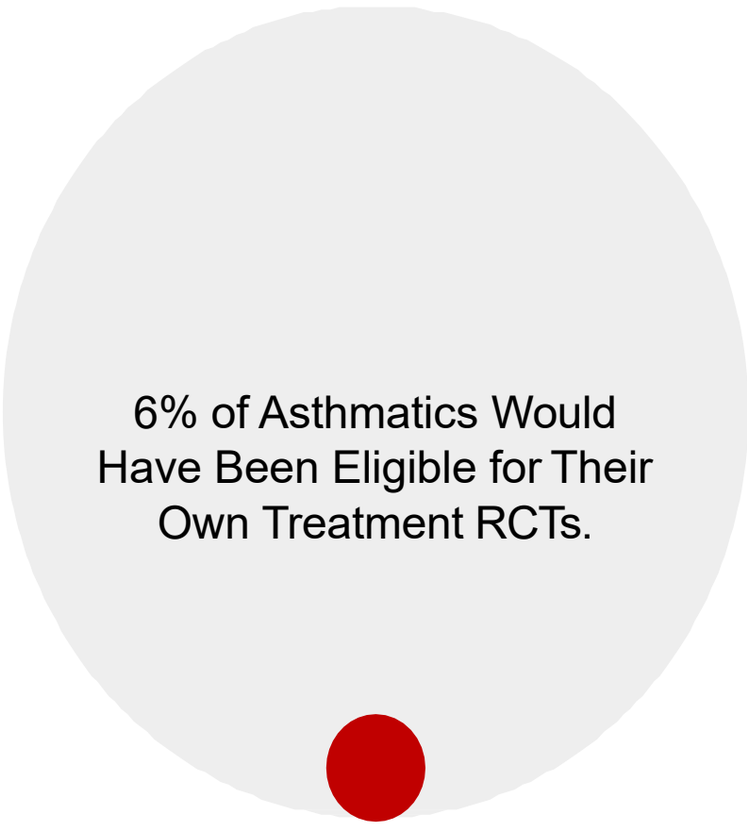
UNIVERSITY OF TORONTO

VECTOR INSTITUTE | INSTITUT VECTEUR

# Evidence in Healthcare and Health?

Randomized Controlled Trials (RCTs) are **rare and expensive**, and can encode **structural biases** that apply to very few people.




*1Smith M, Saunders R, Stuckhardt L, McGinnis JM, Committee on the Learning Health Care System in America, Institute of Medicine. Best Care At Lower Cost: The Path To Continuously Learning Health Care In America. Washington: National Academies Press; 2013.*

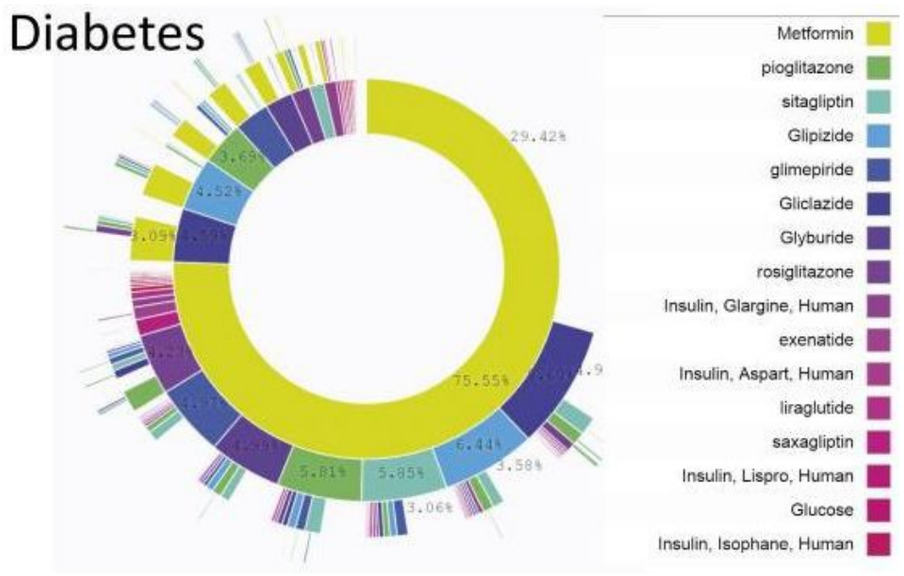
2 Travers, Justin, et al. "External validity of randomised controlled trials in asthma: to whom do the results of the trials apply?." Thorax 62.3 (2007): 219-223.



UNIVERSITY OF TORONTO

VECTOR INSTITUTE | INSTITUT VECTEUR

# Human Treatment Pathways Are Shockingly Unique

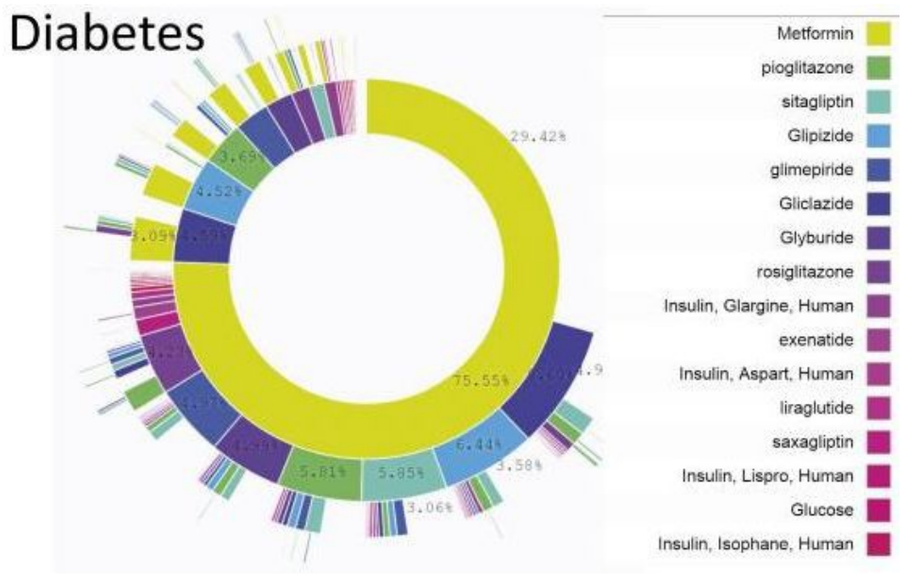



In a combined EHR/claims dataset from 11 sources/4 countries/250 million patients, how many followed a unique treatment pathway?

- Diabetes:
- Depression:
- Hypertension:

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

# Human Treatment Pathways Are Shockingly Unique

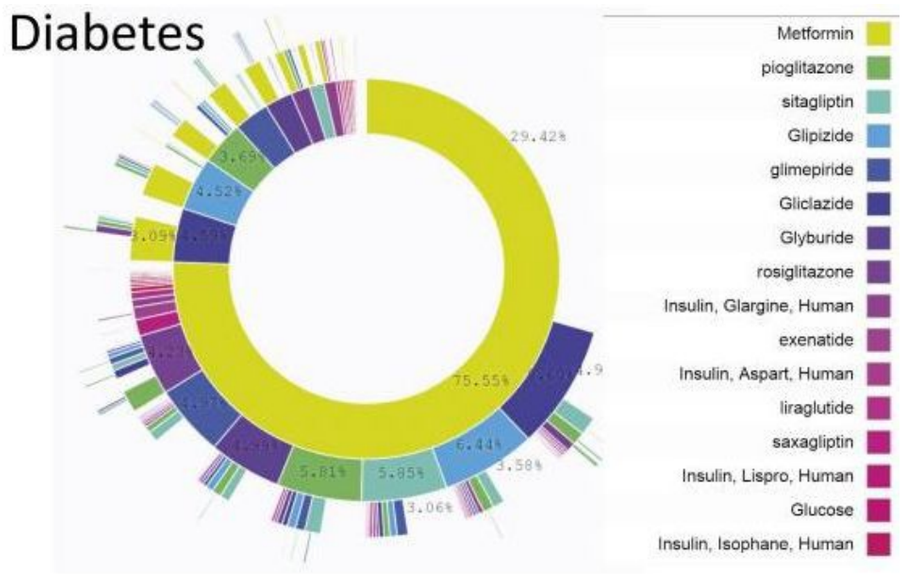



In a combined EHR/claims dataset from 11 sources/4 countries/250 million patients, how many followed a unique treatment pathway?

- Diabetes: **10%** of patients
- Depression:
- Hypertension:

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

# Human Treatment Pathways Are Shockingly Unique

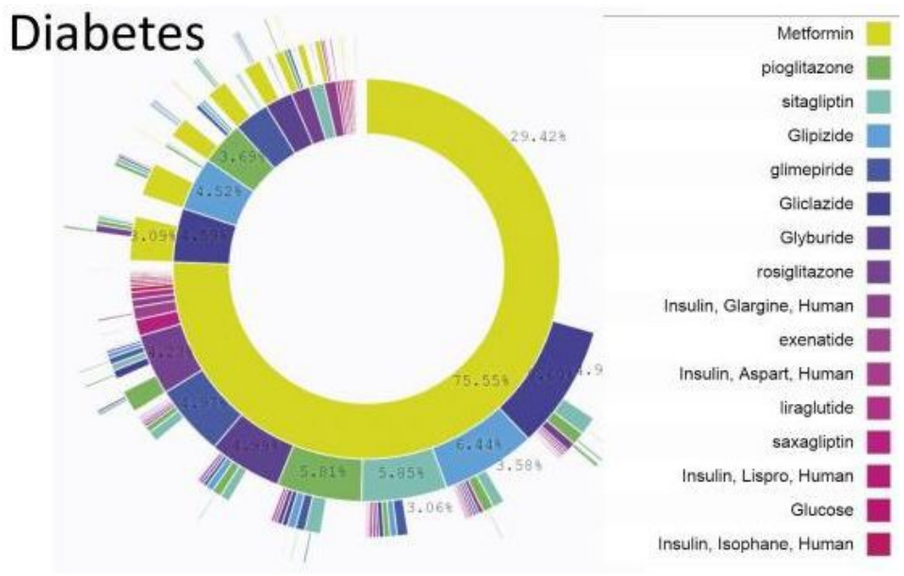



In a combined EHR/claims dataset from 11 sources/4 countries/250 million patients, how many followed a unique treatment pathway?

- Diabetes: **10%** of patients
- Depression: **11%** of patients
- Hypertension:

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

# Human Treatment Pathways Are Shockingly Unique

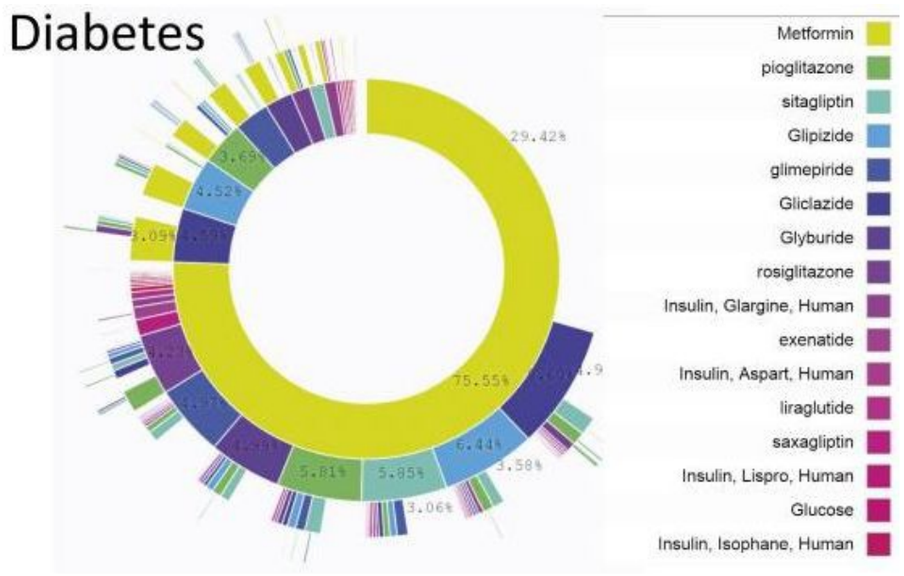



In a combined EHR/claims dataset from 11 sources/4 countries/250 million patients, how many followed a unique treatment pathway?

- Diabetes: **10%** of patients
- Depression: **11%** of patients
- Hypertension: **24%** of patients

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

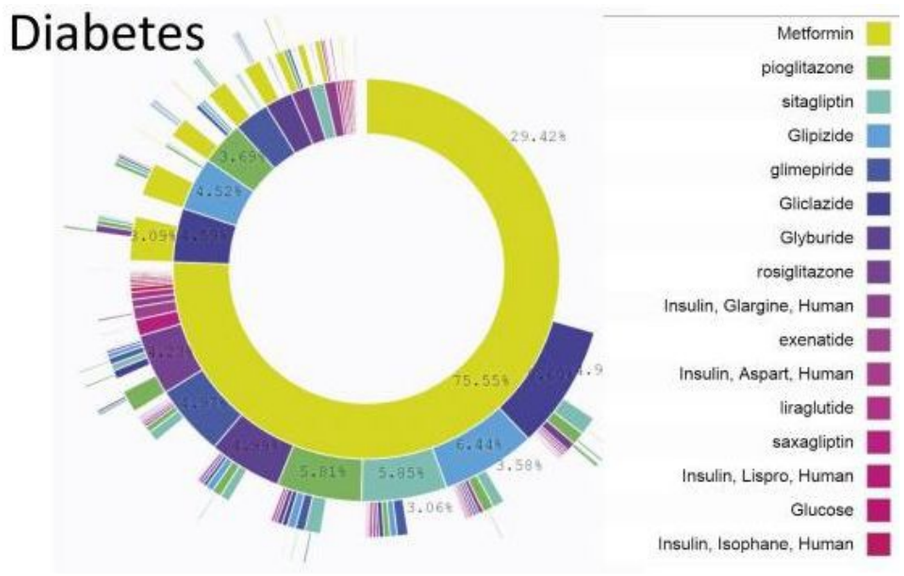
# Human Treatment Pathways Are Shockingly Unique




"In an underlying population of 250 million, based on my 3-y treatment pathway, what patients are like me?"

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

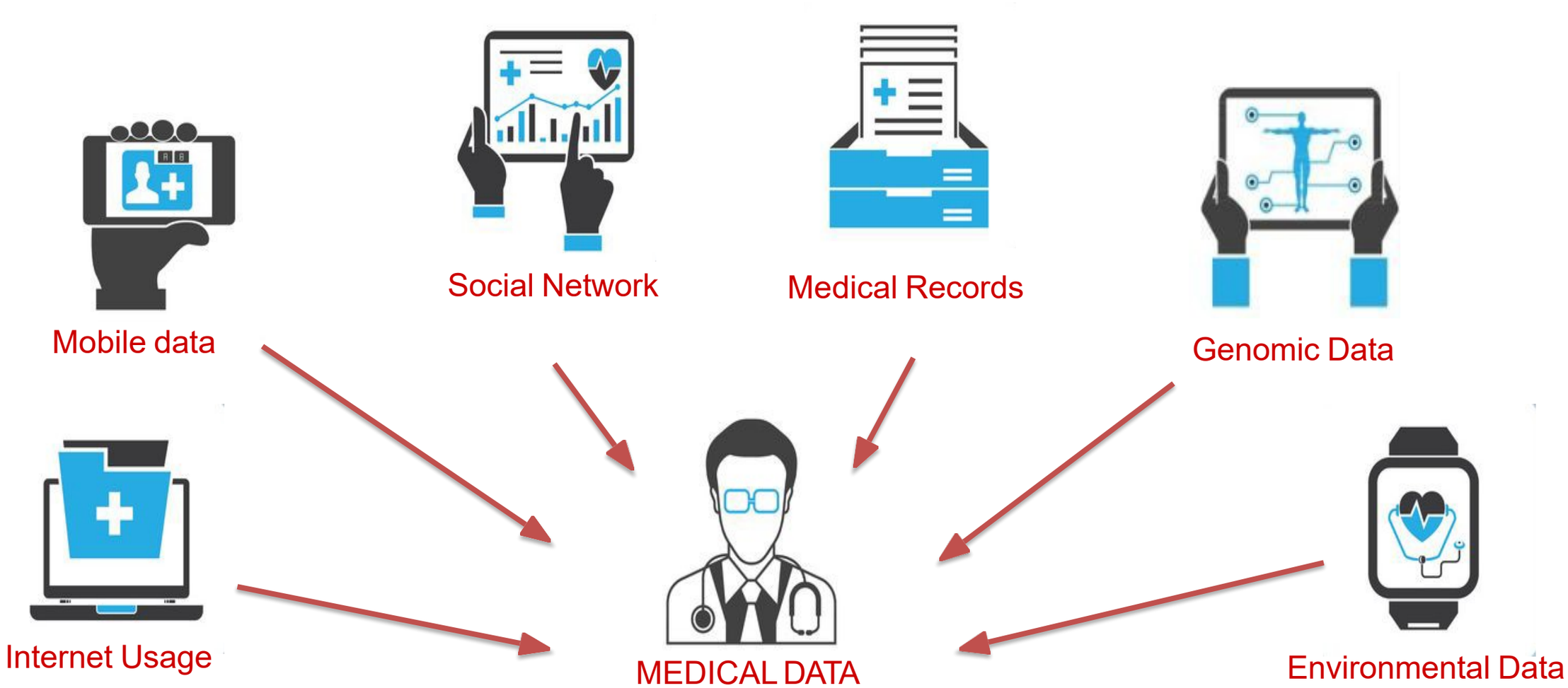
# Human Treatment Pathways Are Shockingly Unique




"In an underlying population of 250 million, based on my 3-y treatment pathway, what patients are like me?"

For 24% of hypertension patients, "**No one.**"

[1] Hripcsak, George, et al. "Characterizing treatment pathways at scale using the OHDSI network." Proceedings of the National Academy of Sciences 113.27 (2016): 7329-7336.

# Opportunity: Machine Learning What Is Healthy

Can we use **data** to **learn** what is **healthy**?