Reinventing the science of medicine with AI

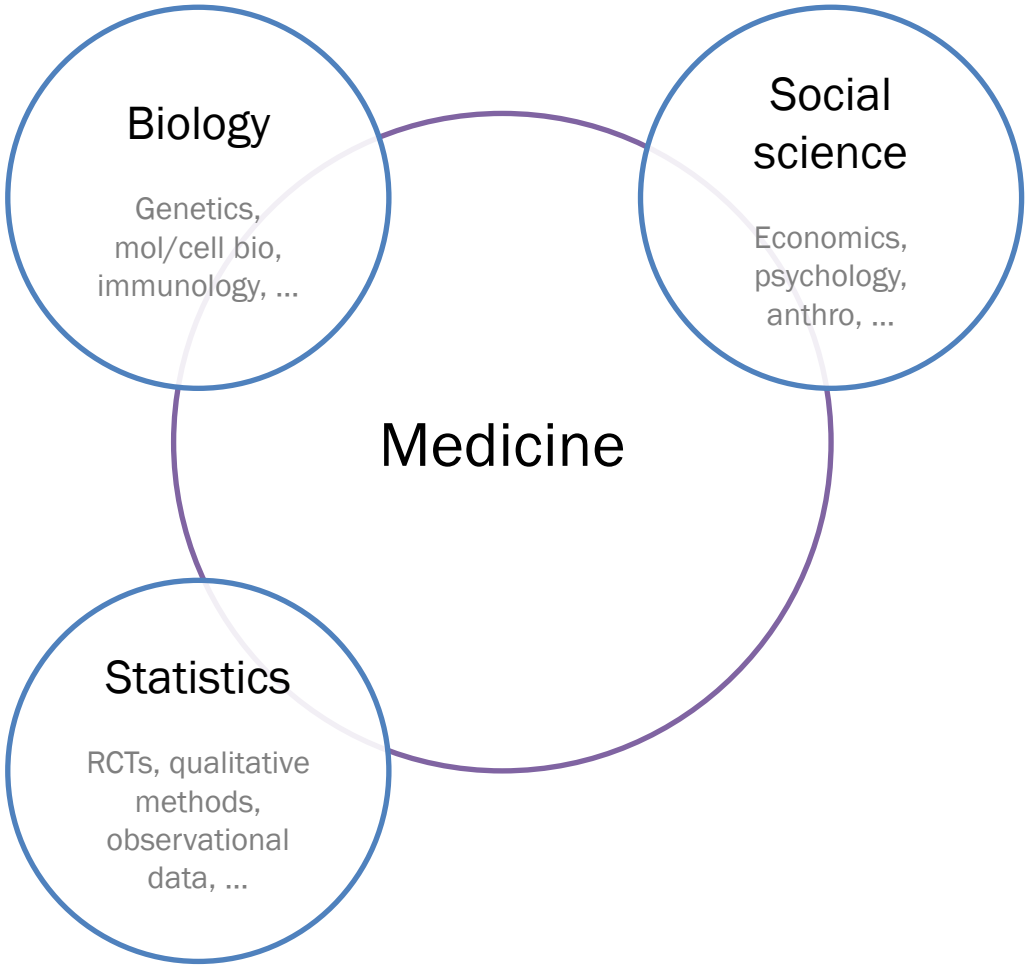
Ziad Obermeyer
UC Berkeley

Today

- Some talks provide answers to important questions
  - This is not that kind of talk
- This talk is about questions
  - New questions we can answer with AI

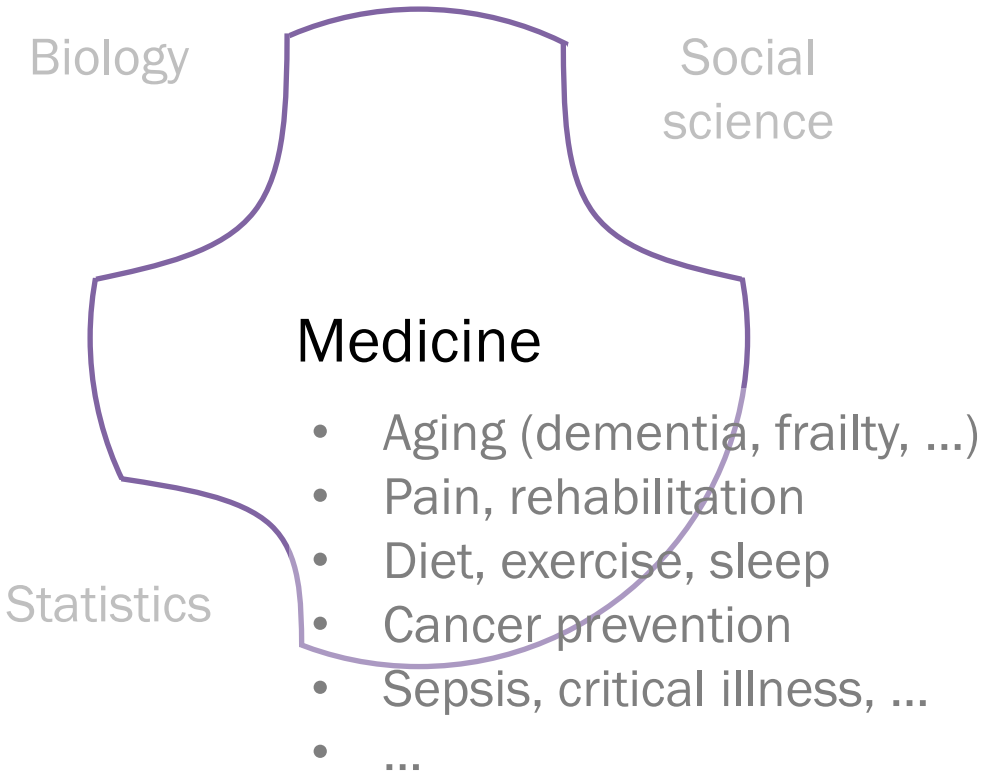
Medical research uses many tools from other fields

- These tools drive major advances
  - CVD
  - Cancer
  - Insurance
  - Drug discovery



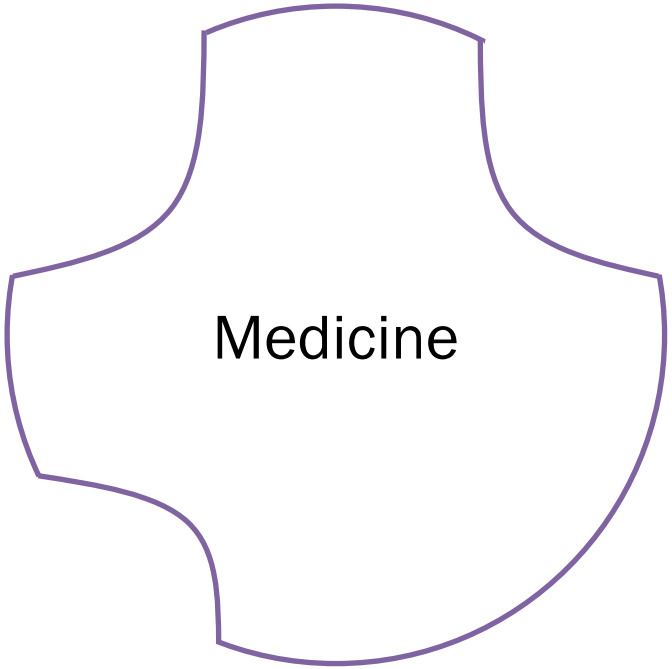
Medical research uses many tools from other fields

- These tools drive major advances
  - CVD
  - Cancer
  - Insurance
  - Drug discovery
- But they also leave huge white space
  - Questions that can't be answered by these tools alone



There are major advances hiding in this white space

- Some historical examples
  - Insulin, cortisone, ALL
  - Cardiac cath, stents
  - Endoscopy, *H. Pylori*
  - Ventilators, ICU
  - Artificial joints

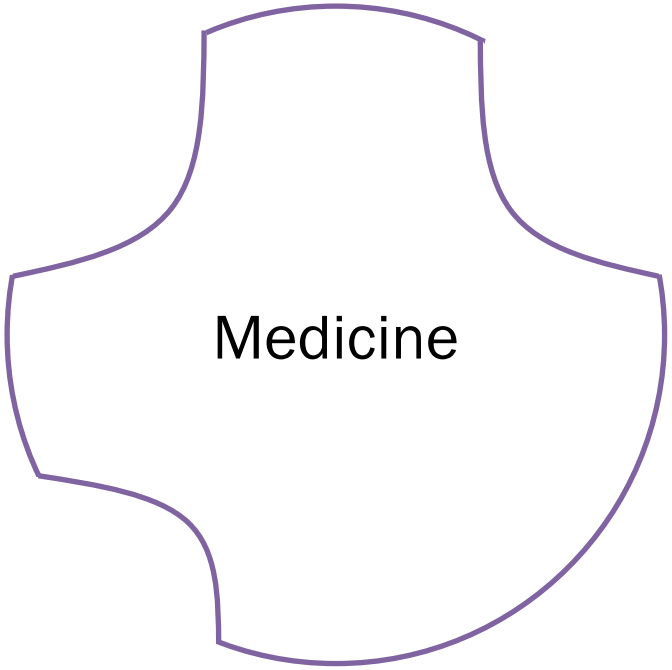


- Secret sauce
  - Painstaking clinical observation
  - Tinkering
  - Trial & error

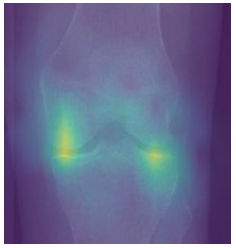
How is this working out today?

AI is a powerful tool for this white space

- Discoveries often start with correlations
  - AI is very good at this
- Humans struggle with complex, high-dimensional data
  - AI is very good at this
- Today: Two [very early] case studies



Two case studies: AI and the new medical science

1. Diagnosis  Knee pain

2. Discovery  Sudden cardiac death

Pain: A major, unsolved clinical problem

- Not just biochemistry, genetics, psychology, …

- Some of the first great medical studies of pain
  - X-ray survey of coal miners vs. office workers
  - Pain correlated with specific radiological findings
  - Later formalized: OA, Kellgren—Lawrence Grade (KLG)

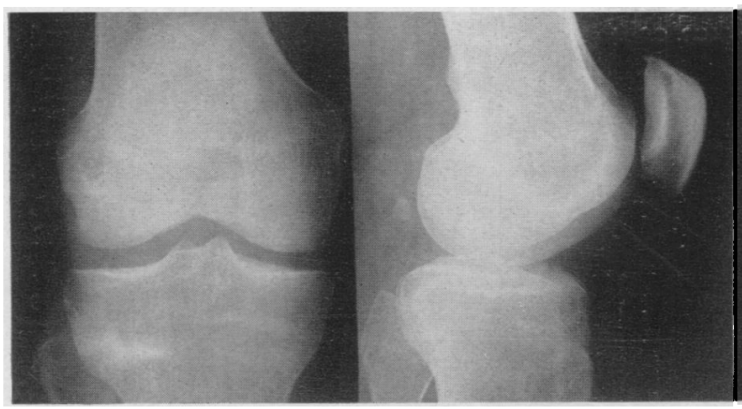


FIG. 5A FIG. 5B

FIGS. 5A and 5B.—Antero-posterior and lateral radiographs of knee showing signs of slight osteo-arthritis.

FIG. 6A FIG. 6B.

FIGS. 6A and 6B..—Antero-posterior and lateral radiographs of knee showing signs of severe osteo-arthritis.

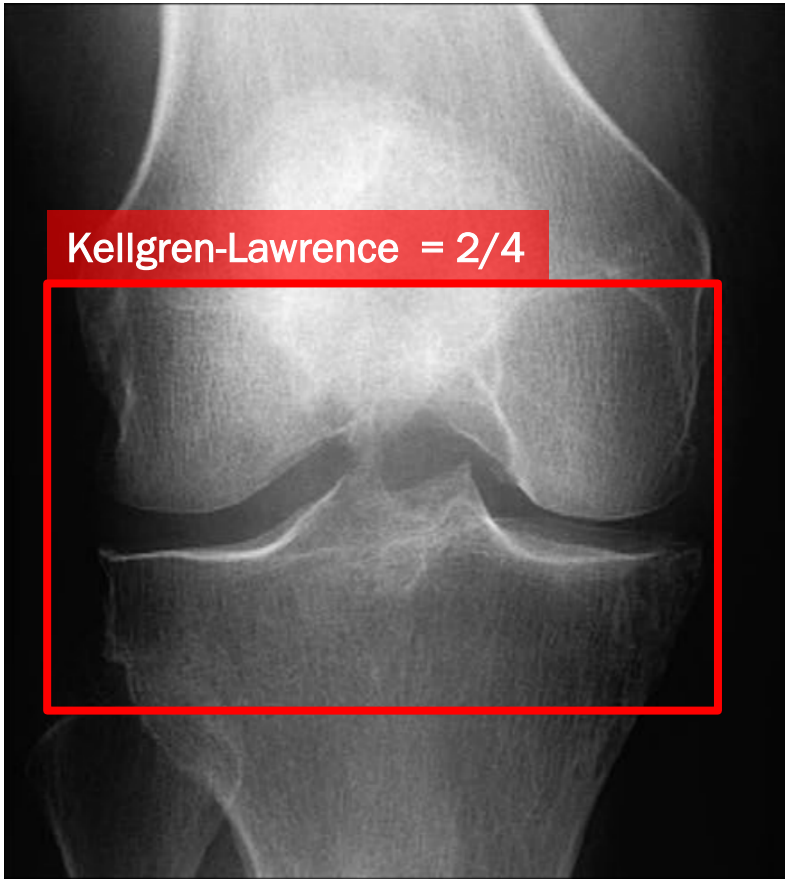
Fast forward: A puzzling "pain gap"—Black vs White patients

- Take patients with same KLG
  - Black patients: more pain
- Partly a social science problem
  - Stress, psychosomatic factors, coping skills
  - Access to therapies
- But could this be a broader 'science of medicine' problem?
  - Is KLG missing something?
  - Would produce "pain despite normal x-ray" phenomenon

Pierson, Cutler, Leskovec, Mullainathan, Obermeyer, *Nature Med* 2021

Can AI help find what KLG is missing?

- Not if AI simply automates radiologist judgment
  - Will replicate same errors, biases as radiologist



RESEARCH ARTICLE Open Access

Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population

Simon Olsson, Ehsan Akbarian, Anna Lind, Ali Sharif Razavian and Max Gordon*

SCIENTIFIC REPORTS

OPEN

Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach

Received: 21 July 2017

Aleksei Tiulpin[1], Jérôme Thevenot[1], Esa Rahtu[3], Petri Lehenkari[2] & Simo Saarakkala[1,4]

Osteoarthritis Data Integration Portal (OsteoDIP): A web-based gene and non-coding RNA expression database

s in Biology and Medicine

Volume 133, June 2021, 104334

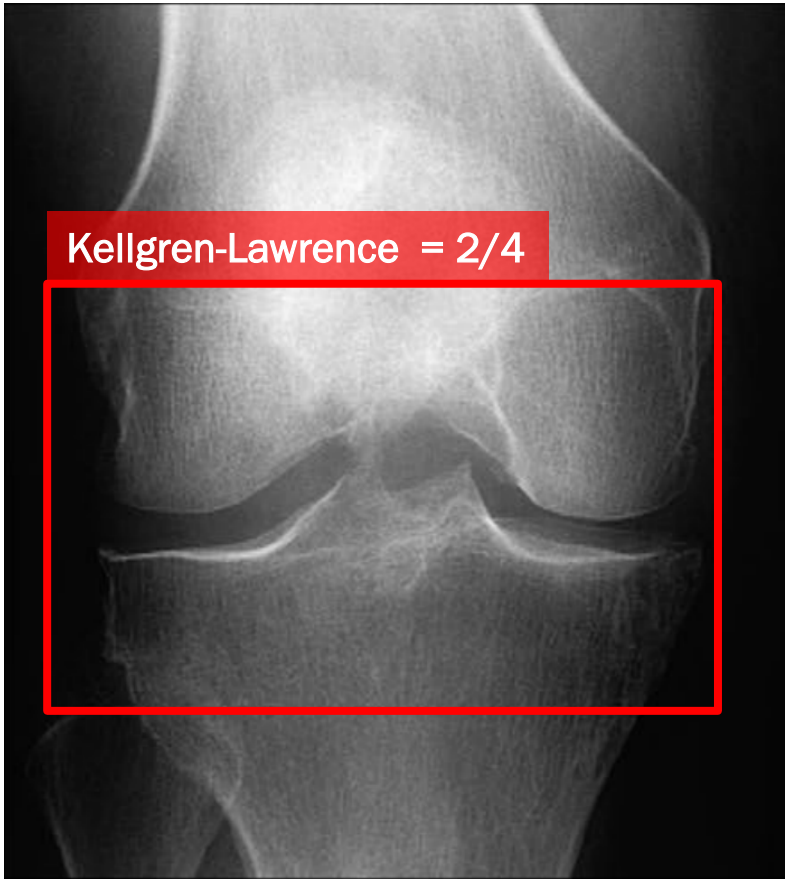
ELSEVIER

Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists

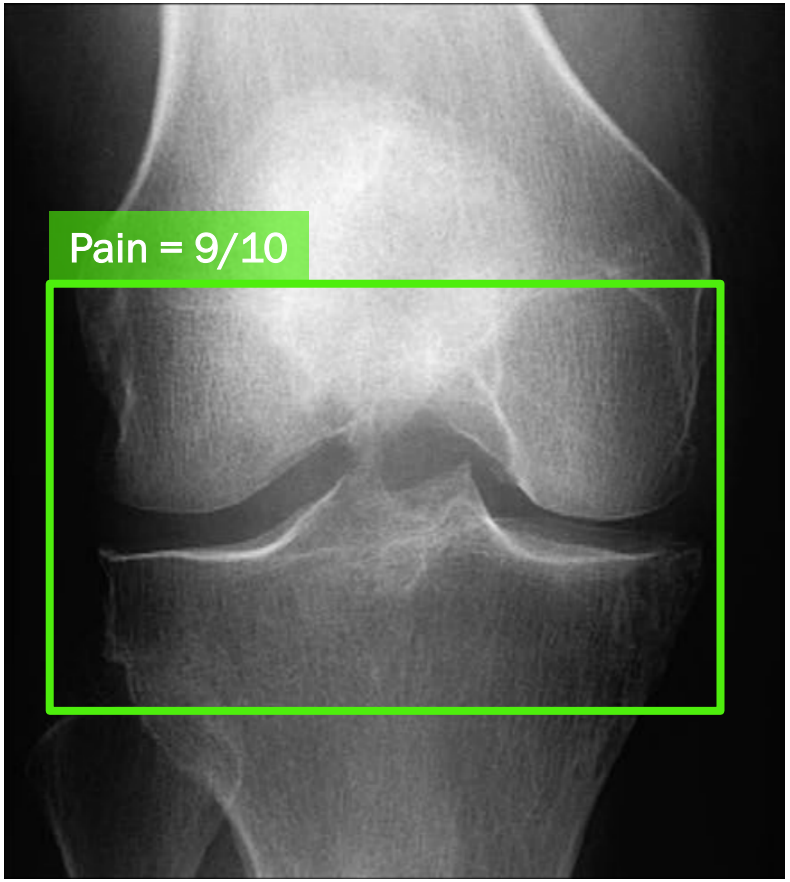
Albert Swiecicki [a], Nianyi Li [a], Jonathan O'Donnell [b], Nicholas Said [b], Jichen Yang [a], Richard C. Mather [c], William A. Jiranek [c], Maciej A. Mazurowski [a b]

A different approach

Learn from the radiologist

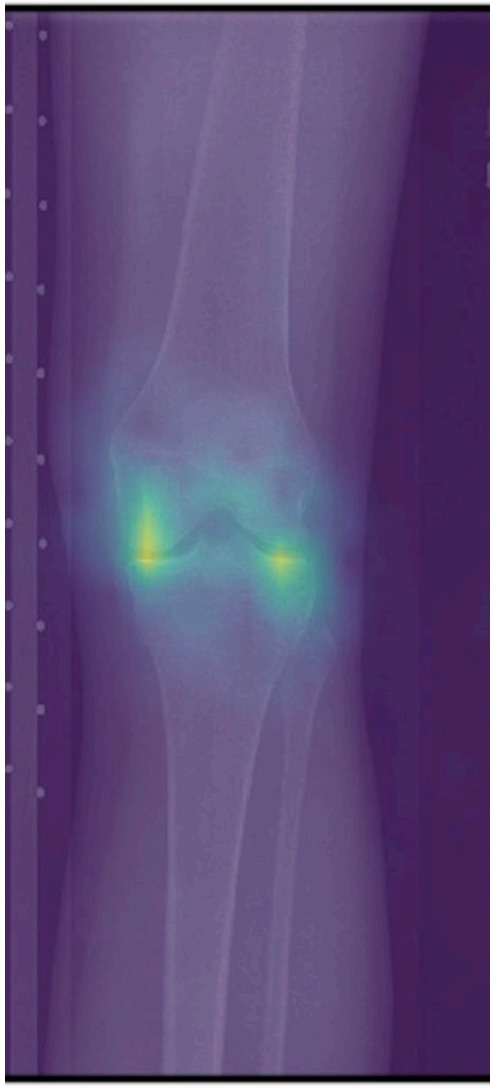


Listen to the patient



Pierson, Cutler, Leskovec, Mullainathan, Obermeyer, *Nature Med* 2021

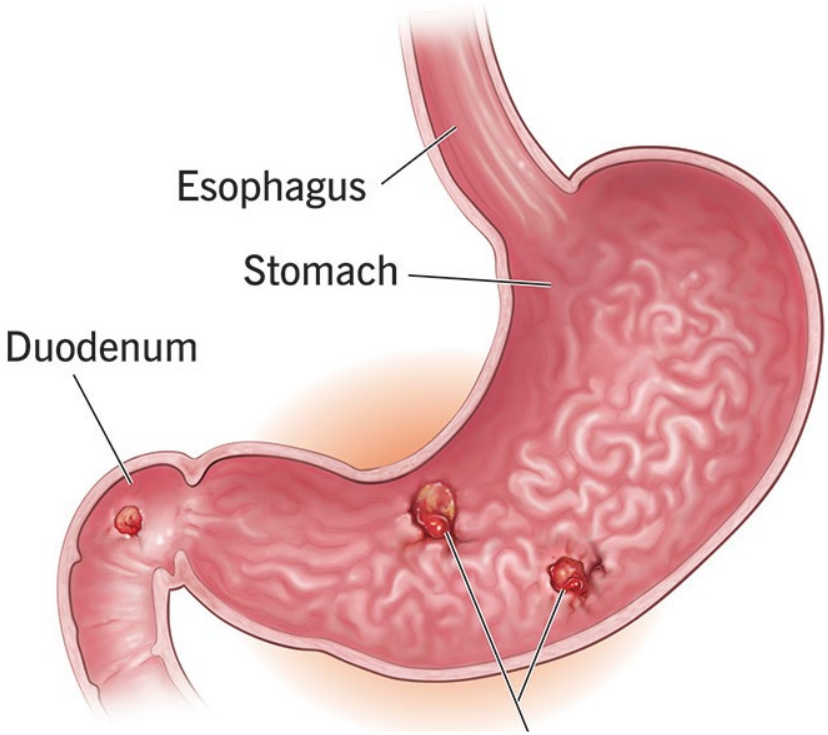
AI explains half of ‘unexplained’ Black-White pain gap



- AI finds new correlates of pain
  - Disproportionately present in Black patients
  - Their knee disease is more severe than we think
- Disparities in knee replacement
  - Insurance doesn’t cover ortho, TKR for normal x-ray
- AI score would double Black knees eligible for TKR

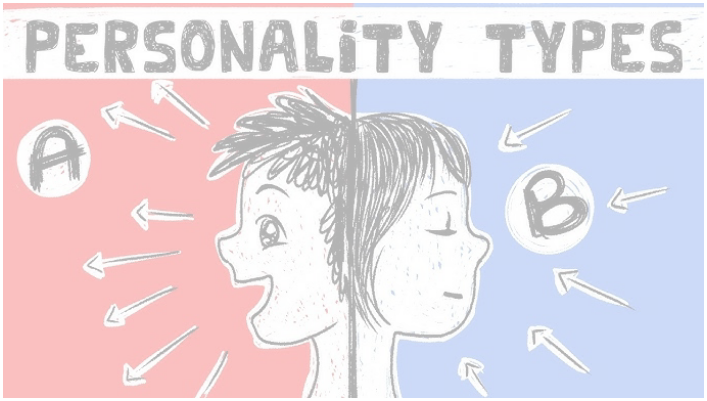
Pierson, Cutler, Leskovec, Mullainathan, Obermeyer, *Nature Med* 2021

Choose the right tools for the job

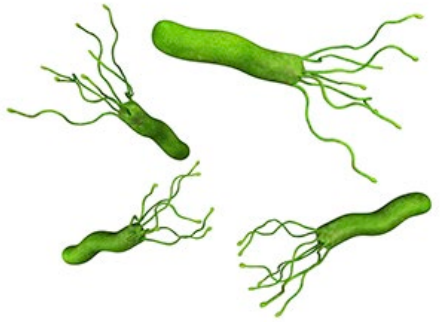


Peptic ulcers

Social science problem?



Medical science problem



Two case studies: AI and the new medical science

1. Diagnosis

Knee pain

2. Discovery

Sudden cardiac death

A medical mystery

- Every year 300-450,000 experience sudden cardiac death
- What makes this even more tragic
  - We have the cure
- We're just very bad at getting the cure to the right patients
  1. False negatives: Many deaths without ICD
  2. False positives: 30-40% of ICDs never fire

Useful to predict who will need this

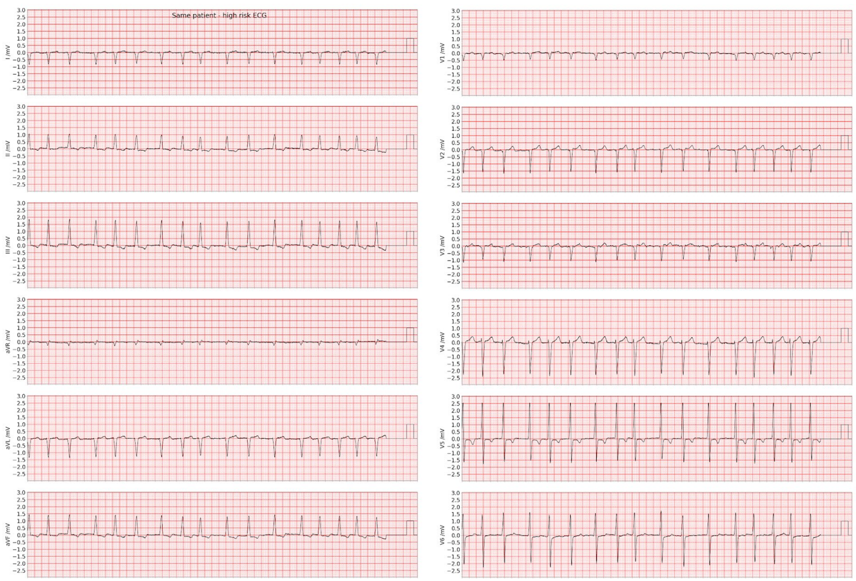
What we do

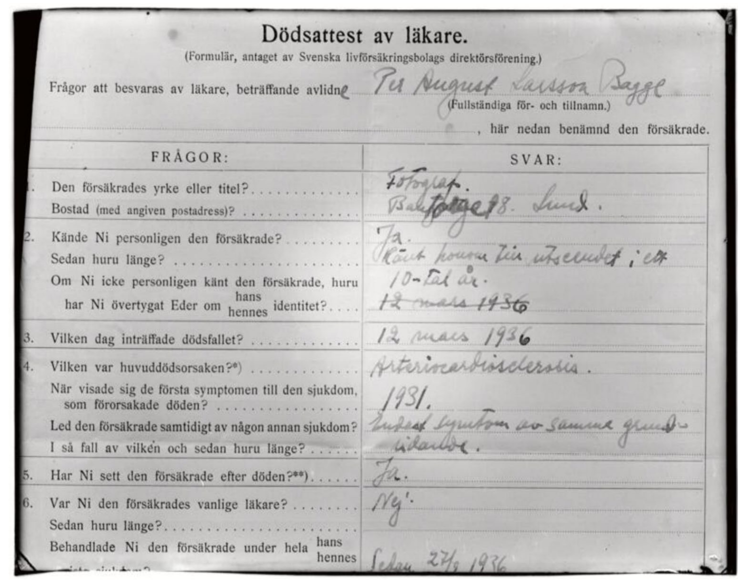
Input: ECG waveform

- All 401,765 ECGs (2014-18)
- From 119,724 patients

Output: Death certificate

- 100% linkage to SCD label
- Full EHR data





Dödsattest av läkare.

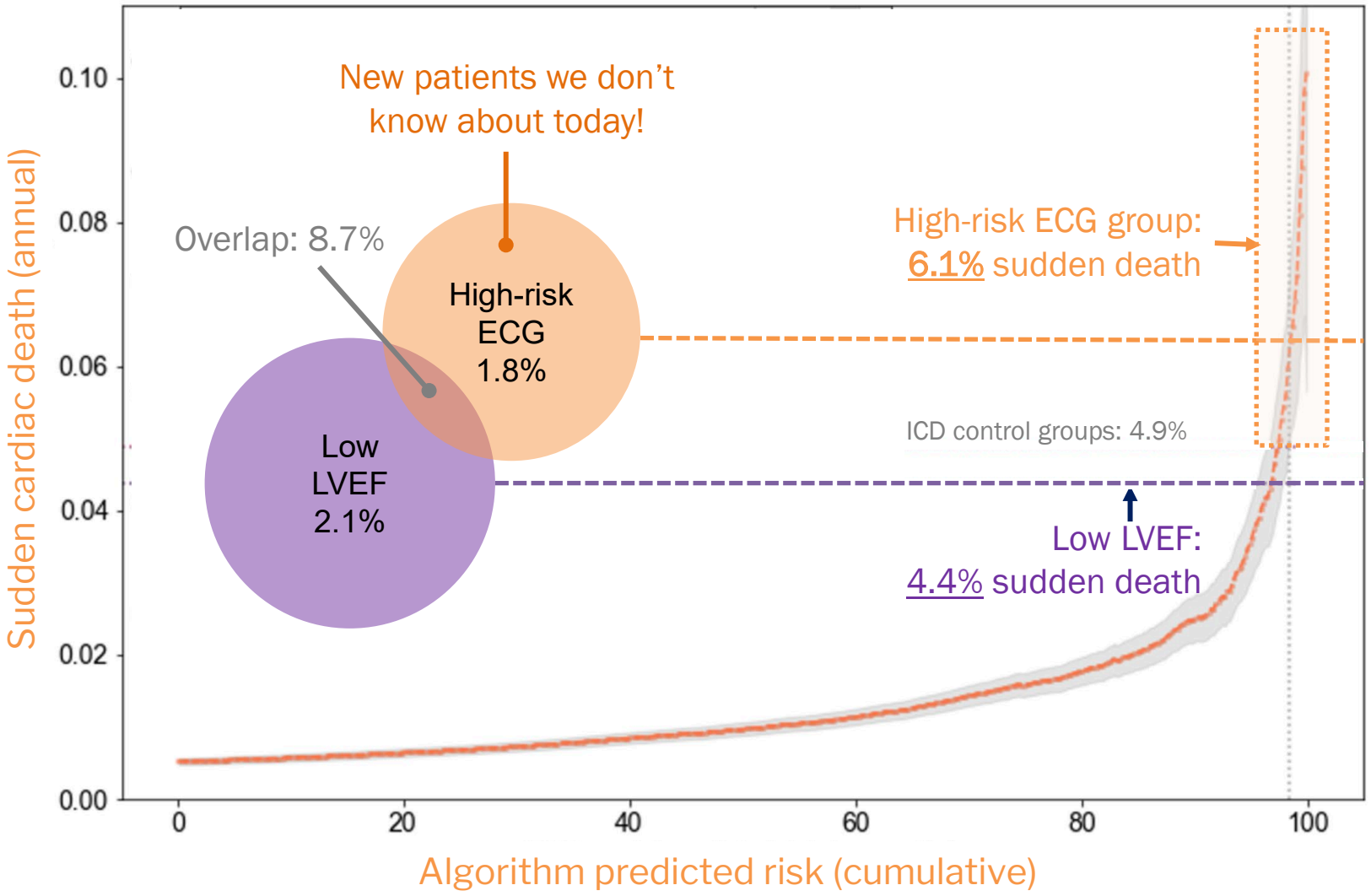
(Formulär, antaget av Svenska livförsäkringsbolags direktörsförening.)

Frågor att besvaras av läkare, beträffande avlidne Per August Larsson Bagge
(Fullständiga för- och tillnamn.)
, här nedan benämnd den försäkrade.

FRÅGOR:	SVAR:
1. Den försäkrades yrke eller titel?	Fotograf.
Bostad (med angiven postadress)?	[illegible] 18. Lund.
2. Kände Ni personligen den försäkrade?	Ja.
Sedan huru länge?	[illegible]; ett 10-tal år.
Om Ni icke personligen känt den försäkrade, huru har Ni övertygat Eder om hans/hennes identitet?	~~12 mars 1936~~
3. Vilken dag inträffade dödsfallet?	12 mars 1936
4. Vilken var huvuddödsorsaken?*)	Arteriocardiosclerosis.
När visade sig de första symptomen till den sjukdom, som förorsakade döden?	1931.
Led den försäkrade samtidigt av någon annan sjukdom?	[illegible]
I så fall av vilken och sedan huru länge?	
5. Har Ni sett den försäkrade efter döden?**)	Ja.
6. Var Ni den försäkrades vanlige läkare?	Nej.
Sedan huru länge?	
Behandlade Ni den försäkrade under hela hans/hennes	[illegible] 1936

Obermeyer et al., *work in progress*

ECG finds new, unsuspected high-risk patients



Obermeyer et al., *work in progress*

Deaths are predictable—but are they preventable?

- High-risk patients with ICDs die 56% less than predicted
  - Sanity check: low LVEF patients with ICDs die 53% less
  - ECG high risk only: Similar reductions all-cause mortality

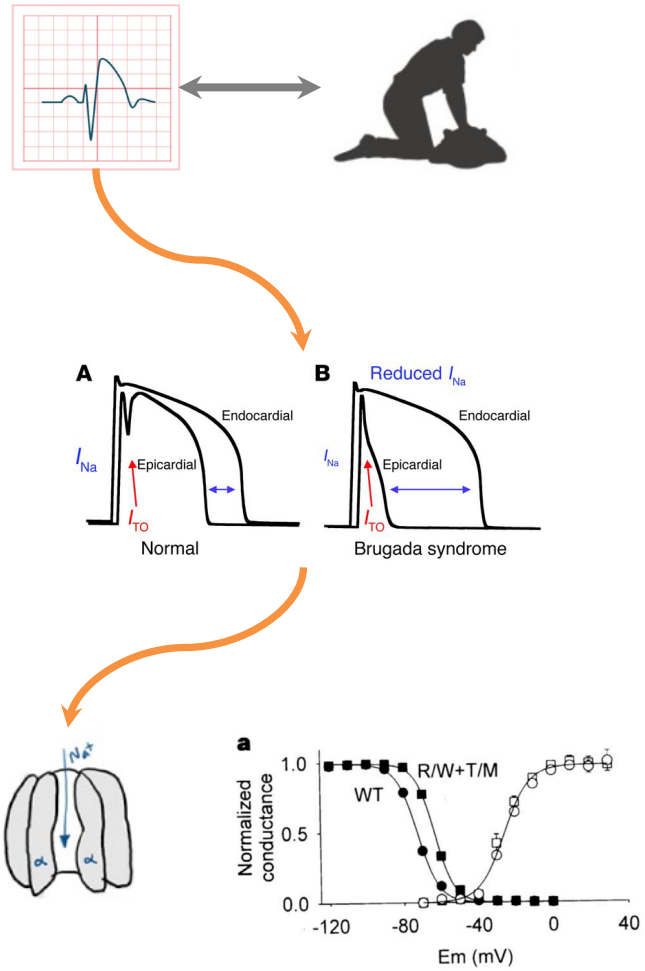
- Taiwan registry: ED arrests & matched controls
  - Take ECGs done >1yr before ED visits, predict risk
  - Model discriminates primary cardiac arrests vs controls
    - Does not discriminate non-cardiac arrest vs controls

- Southern California EHR cohort: 251,858 ECGs (2021-22)
  - High-risk: 19.1% per year have VF/VT (AUC: 0.792)

Obermeyer et al., *work in progress*

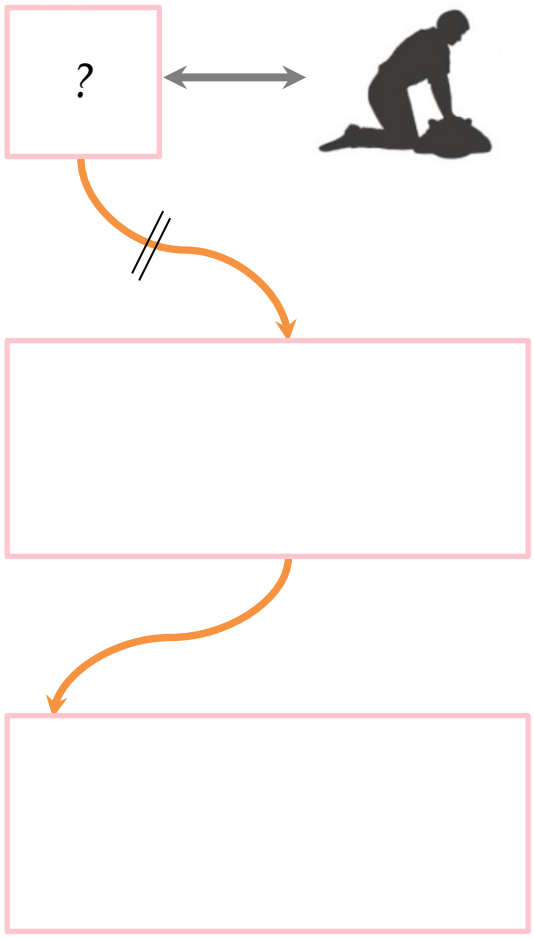
Typical ‘playbook’ for discoveries (made by humans)

1. Notice correlation
   - Show it’s robust
2. Reason about cause
   - Using first principles
3. Test hypotheses
   - Elegant experiments

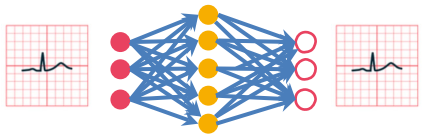


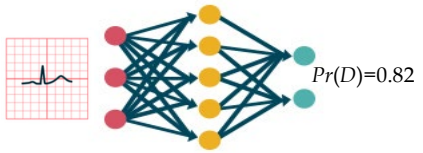
A problem with ‘discoveries’ made by AI

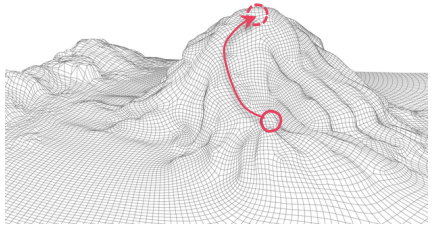
1. Show robust correlation
   - ECGs are high risk
   - …But no waveform!
2. Can’t reason about cause
3. Can’t test hypotheses



A way to visualize what the model is ‘seeing’

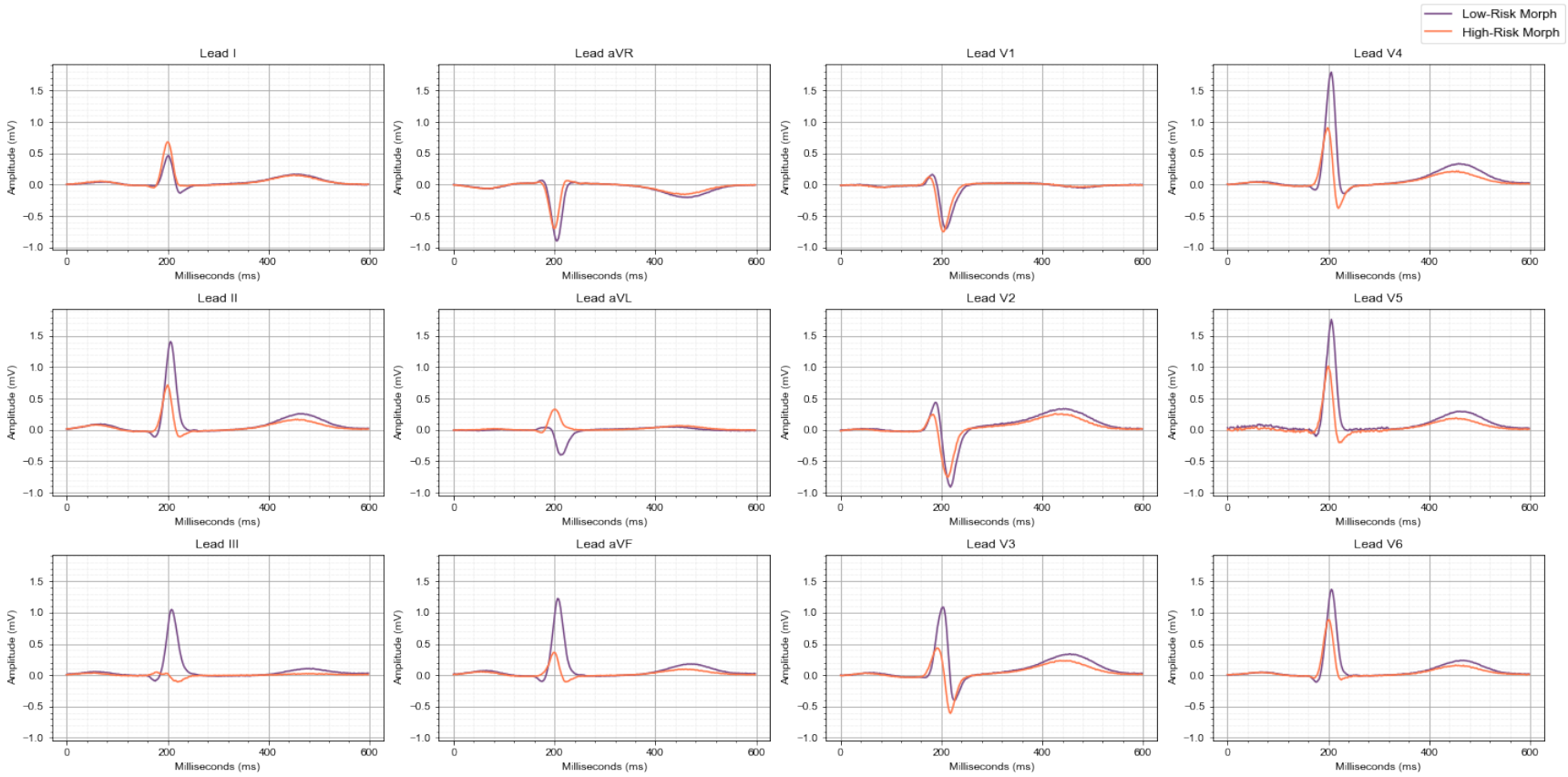




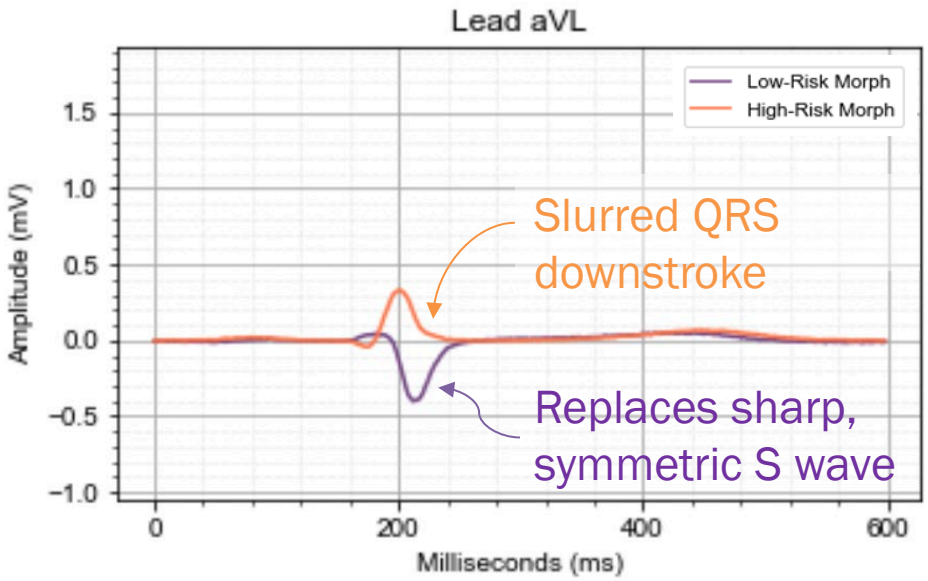


- Train a generative model
  - Encode patients’ ECGs
- ‘Coach’ it with predictive model
  - Calculate risk gradient
- ‘Morph’ a patient’s ECG
  - Follow risk gradient
  - Generate new synthetic ECG
  - …Repeat

Result: A representative morph



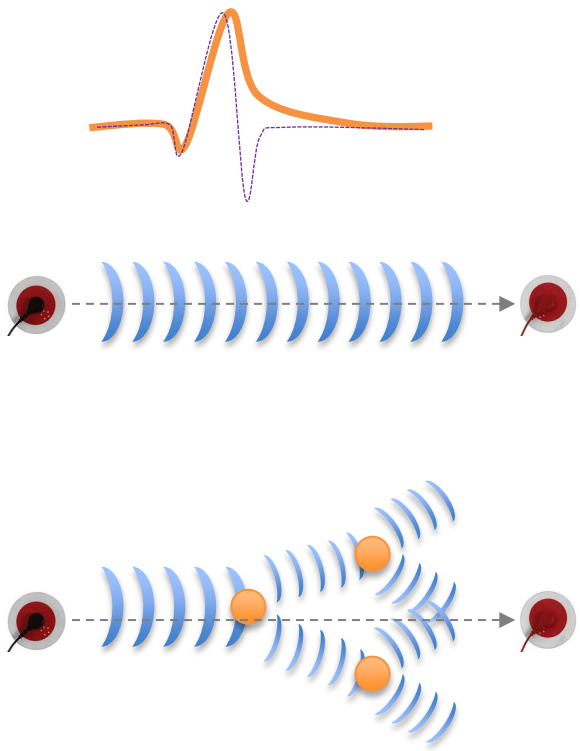
An intriguing feature of high-risk morphs



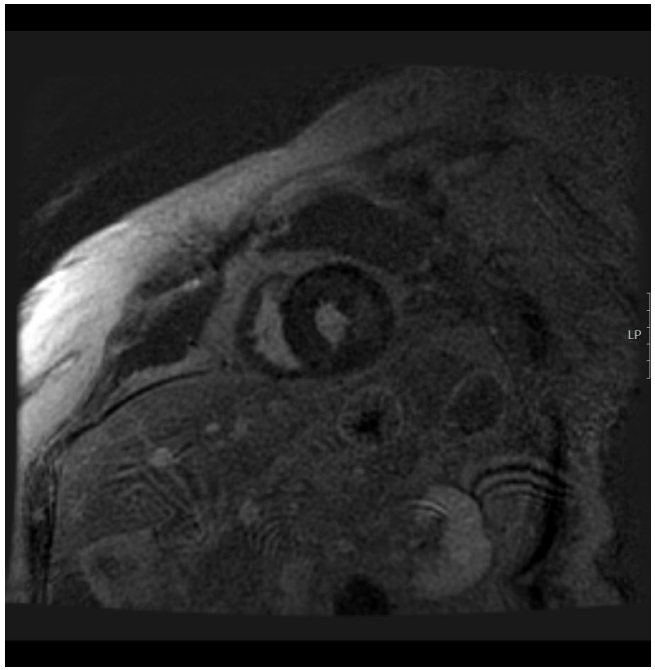
- Qualitative insight
  - Slope decreases over QRS duration
- Quantitative features
  - 1st and 2nd differences
- New features predict sudden death, VF/VT
  - In Sweden
  - In Taiwan, California

Tying empirical observation to physiology

1. Hypothesis generation: Why does QRS flatten?
   - A simple model: scattering



2. Data collection to test hypotheses: cMRI
   - High-risk: 43% LGE
   - Low risk: 7%



Summary: AI and the new science of medicine

- AI is not just another tool for medical research
  - Is will be central to a new science of medicine
- AI will dramatically improve existing science
  - Better decision rules (>5 variables, personalized)
  - Better qualitative methods (summarization, chat)
  - Better RCTs (predictive enrollment, imaging surrogates)
- It will also open up whole new areas of study
  - Making sense of high-dimensional data (diagnosis)
  - Capturing complex relationships (disease models)