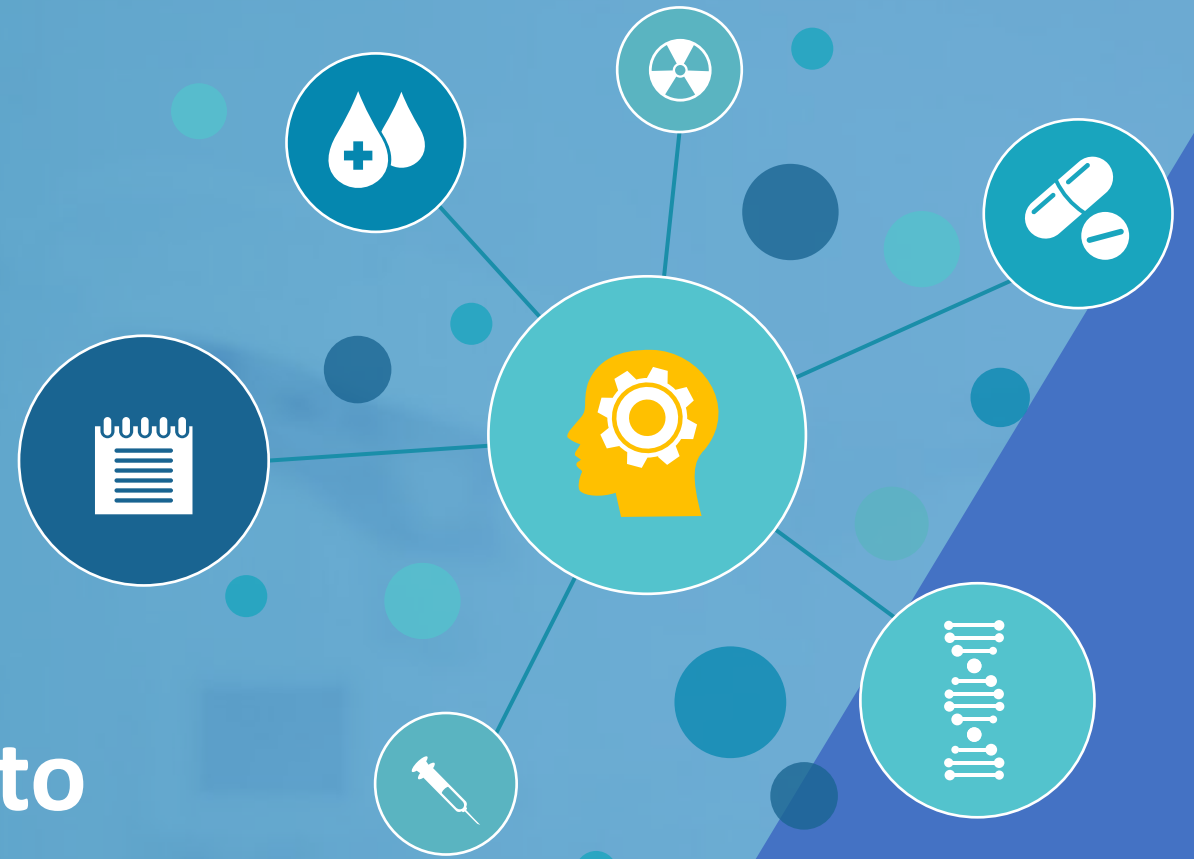


# May AI replace classic statistical approaches to predict clinical outcomes?

Jan Poleszczuk

Department of Computational Oncology

Maria Skłodowska-Curie National Research Institute of Oncology, Poland



# Relevant Financial Disclosure(s)

- CliniNote Company
  - shareholder
  - Chief Technology Officer (wages)



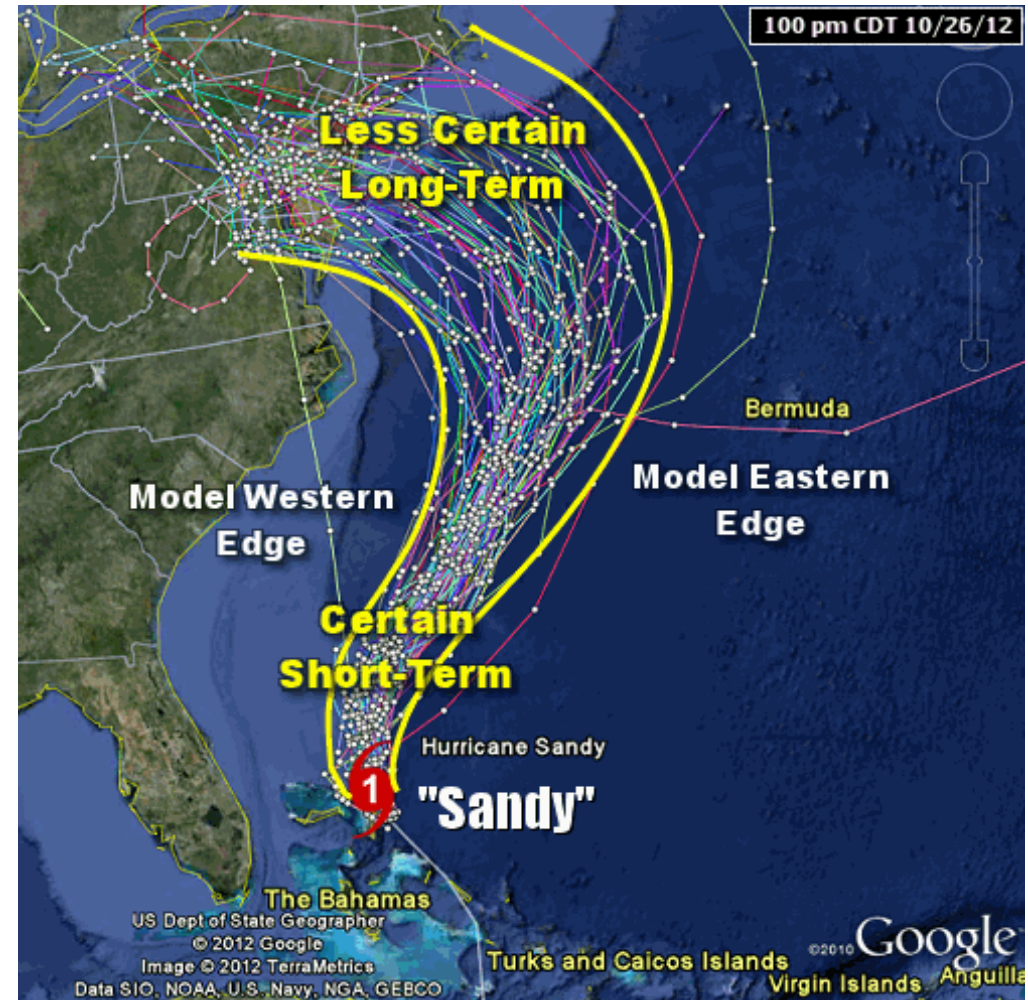
# Clinical outcomes prediction

## Goal

Dynamic prediction of patient response with the smallest possible *cone of uncertainty*.

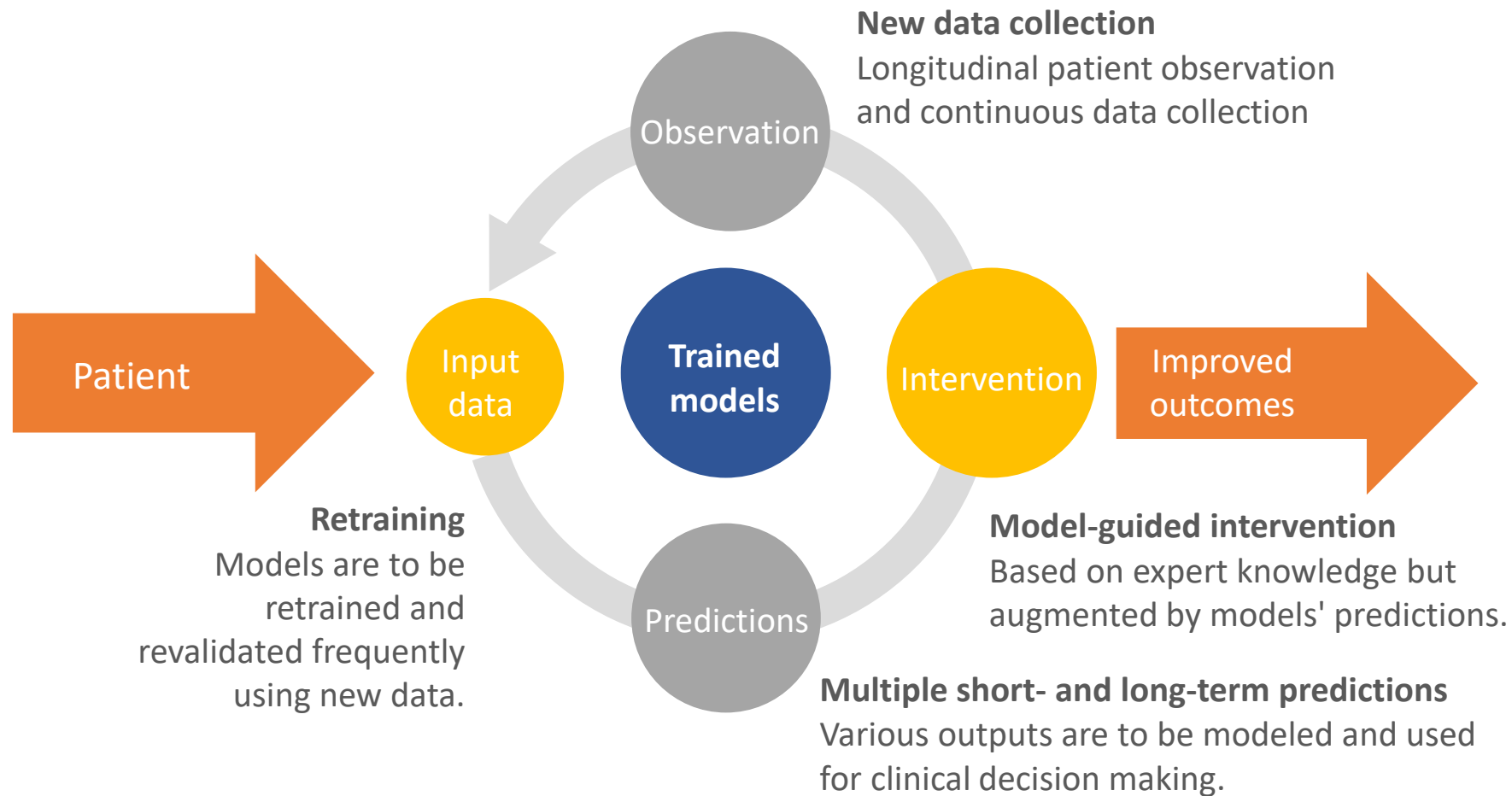
Frequent prediction updates to limit impact of long-term uncertainty.

Actionable predictions are the most important ones.

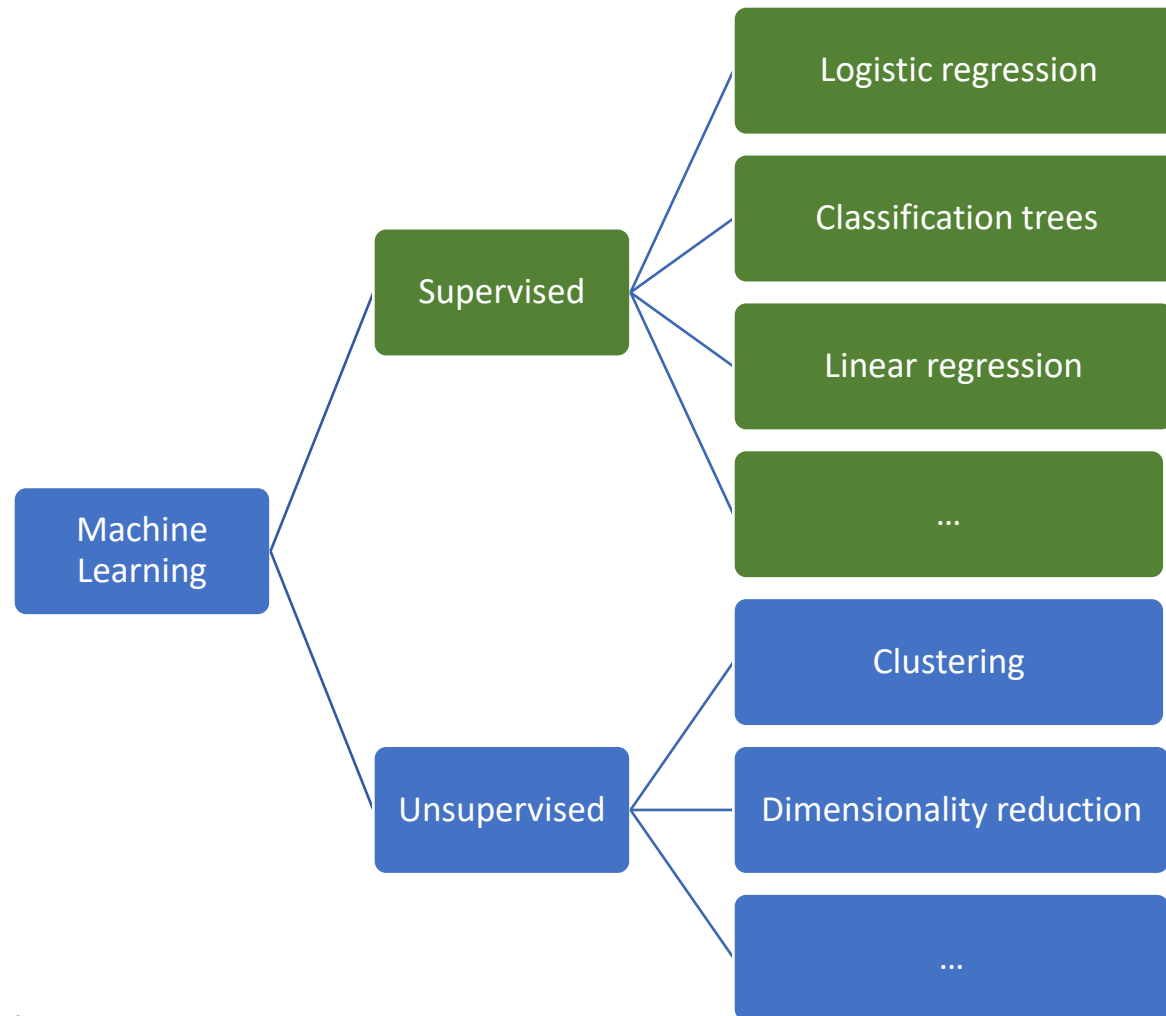


US Dept. Of State Geographer, ©2012 Google, Image ©2012 TerraMetrics, Data SIO, NOAA, U.S. Navy, NGA, GEBCO

# Workflow

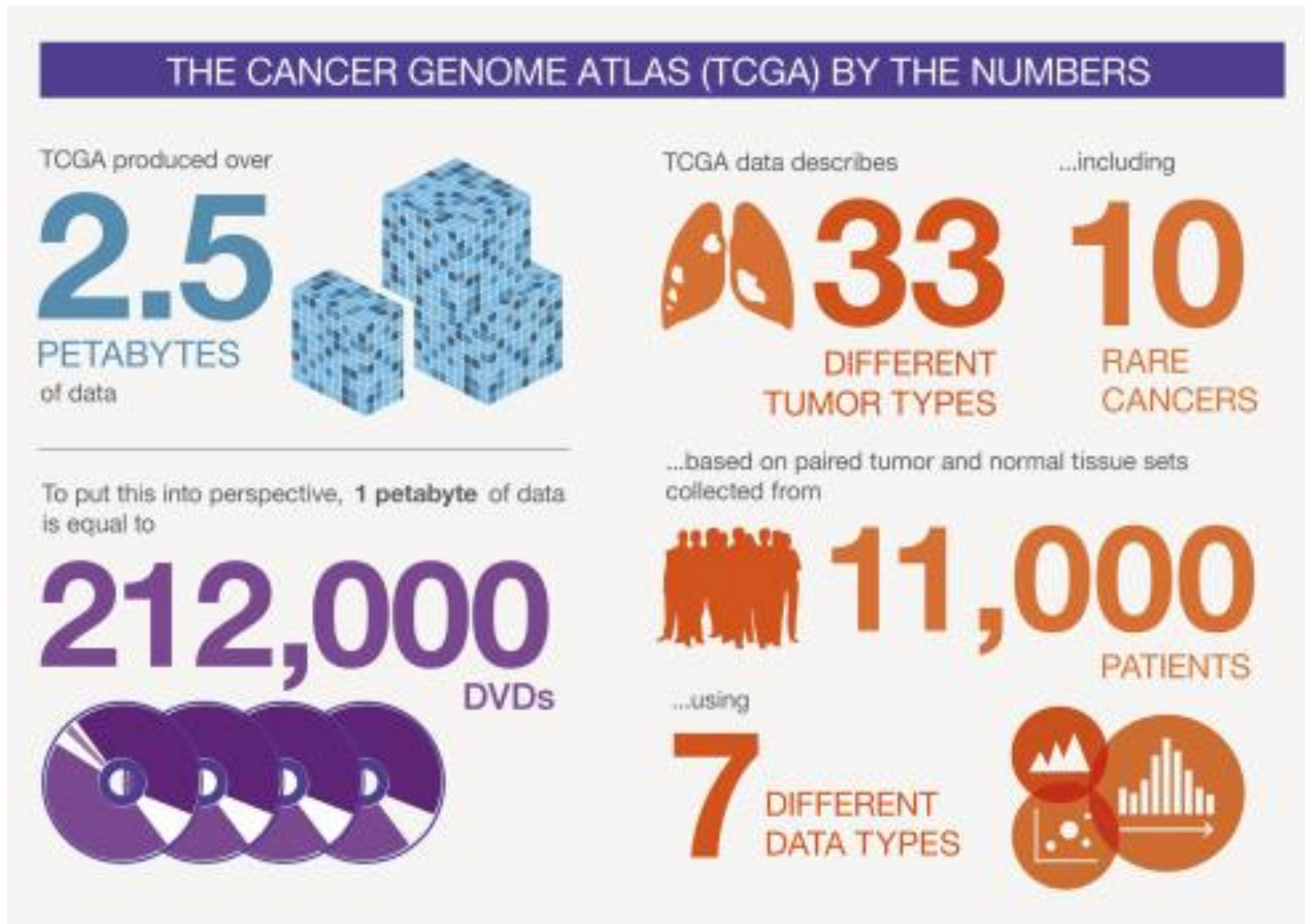


# Creating models – classic approaches



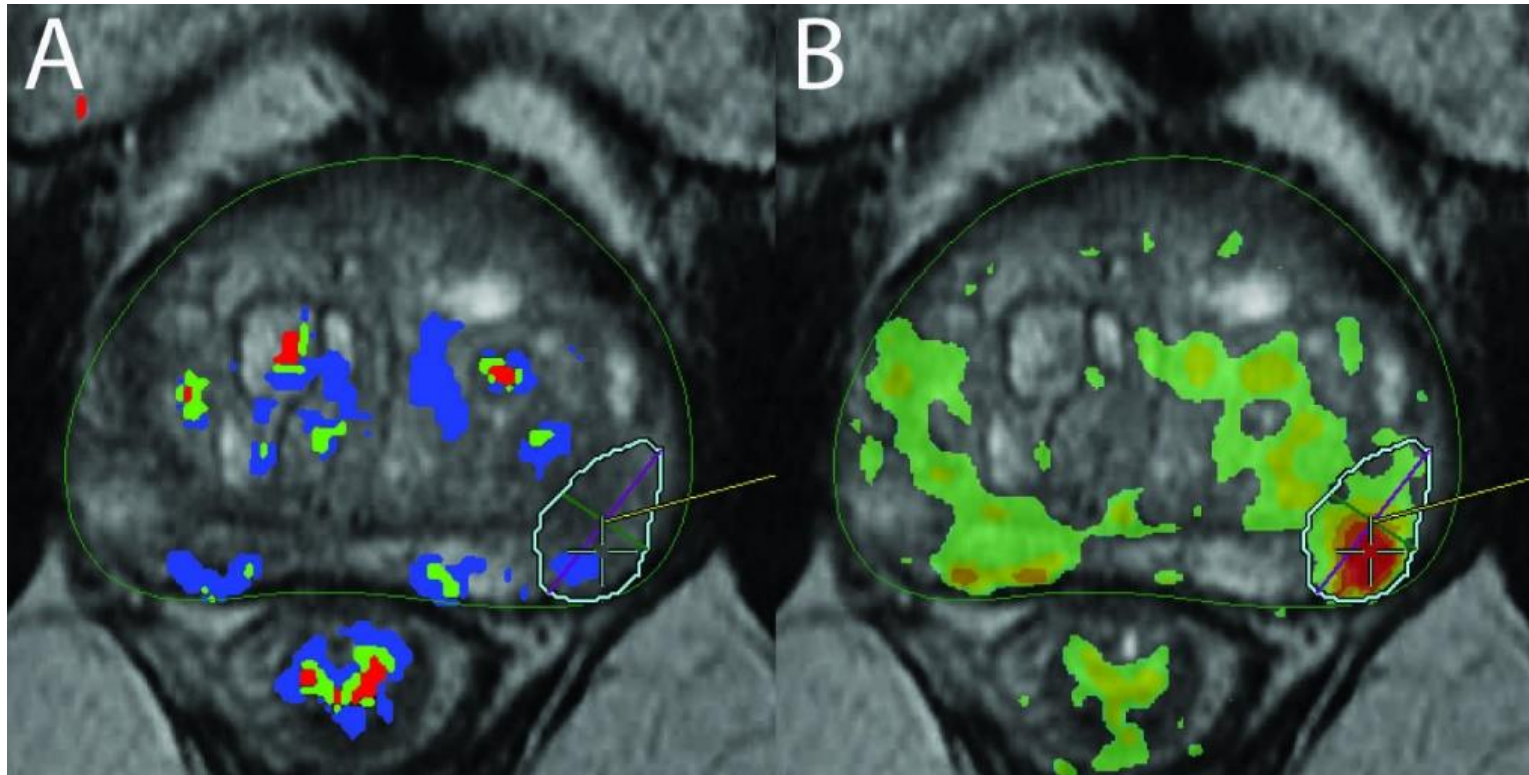
Highly augmented by expert knowledge.

# Huge data volumes



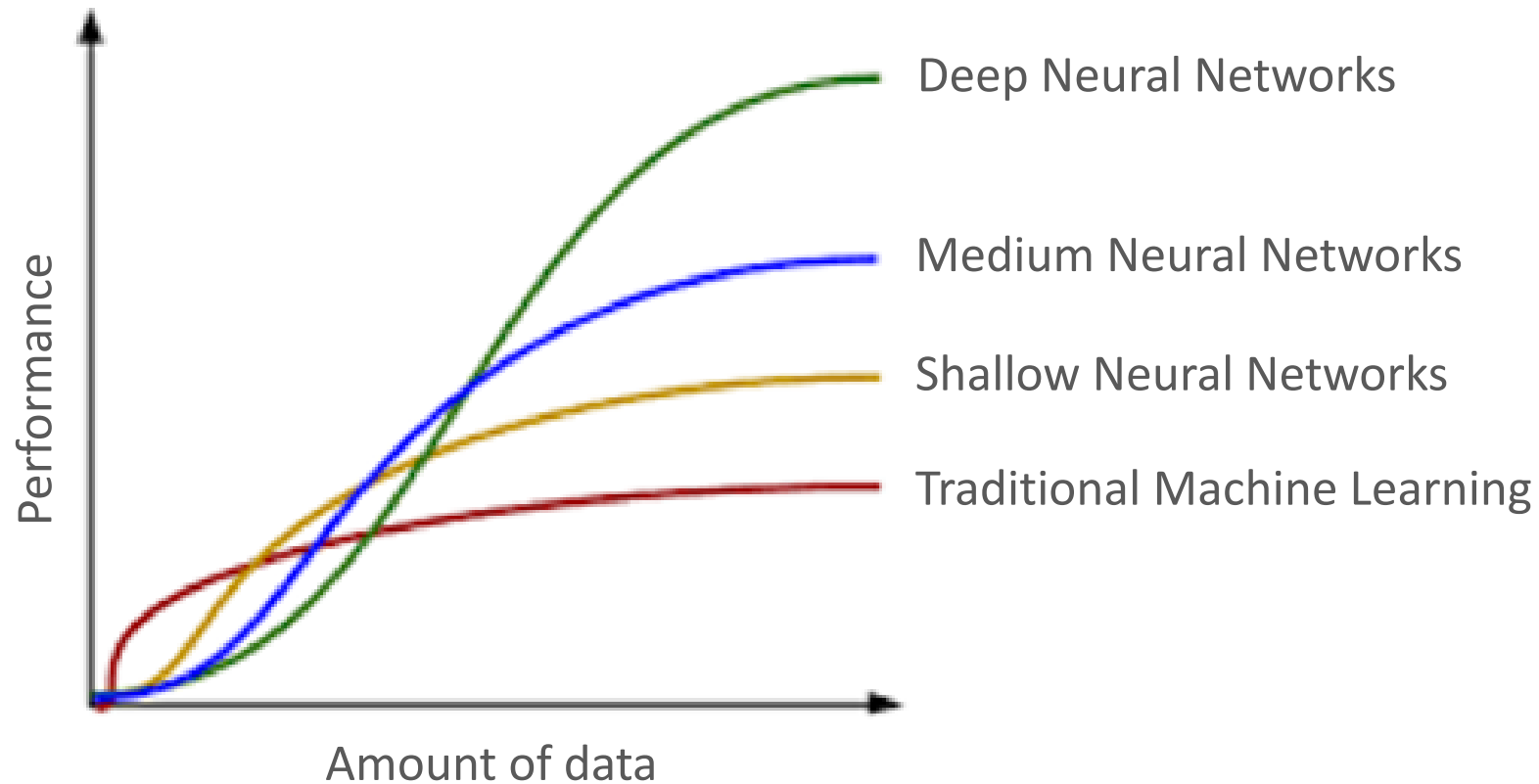
# New data sources and techniques

Functional imaging, proteomics, single cell profiling, histochemistry, etc...



<https://www.nibib.nih.gov/>

# Performance of the models

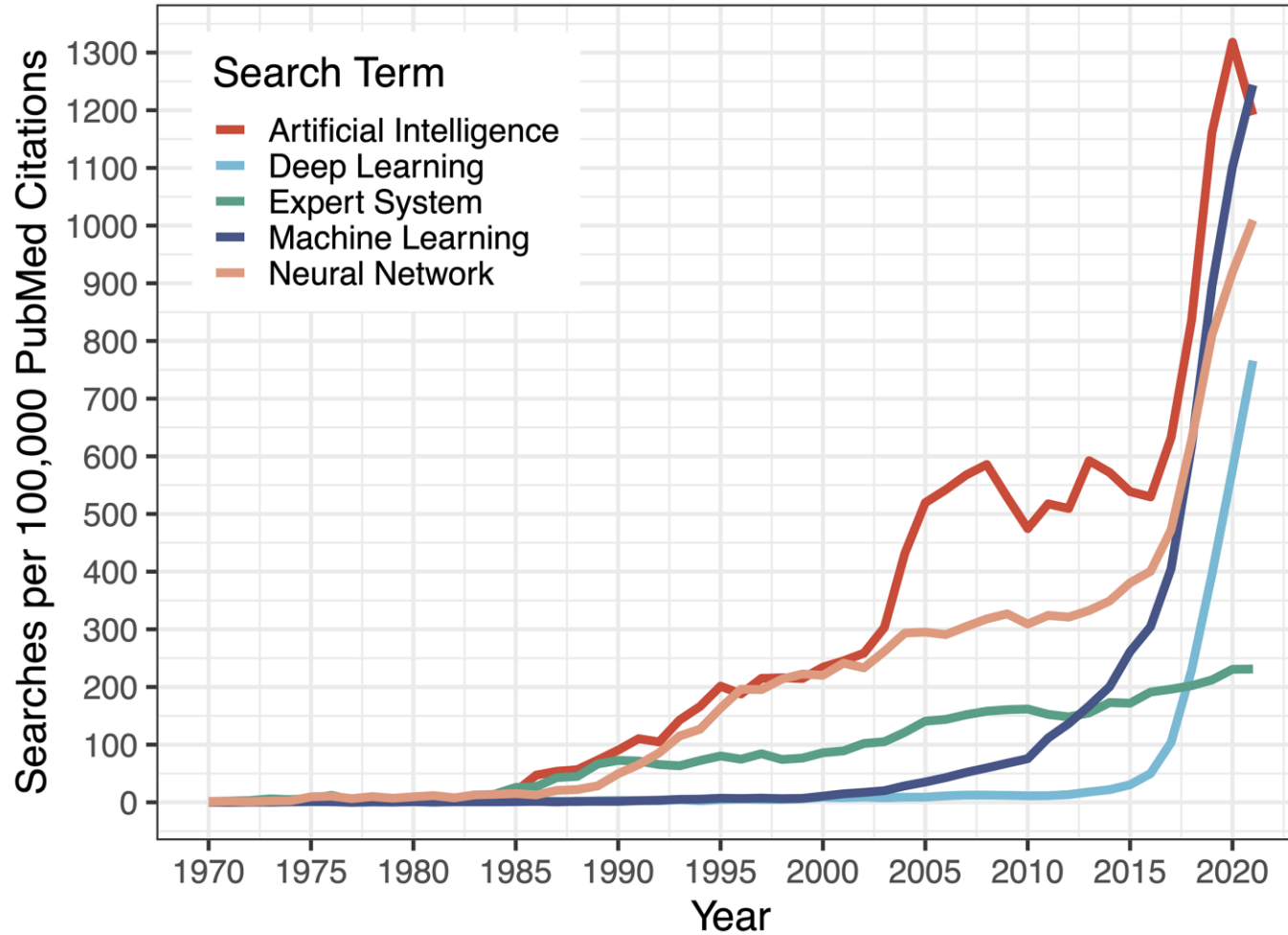


<https://towardsdatascience.com/deep-learning-in-science-fd614bb3f3ce>

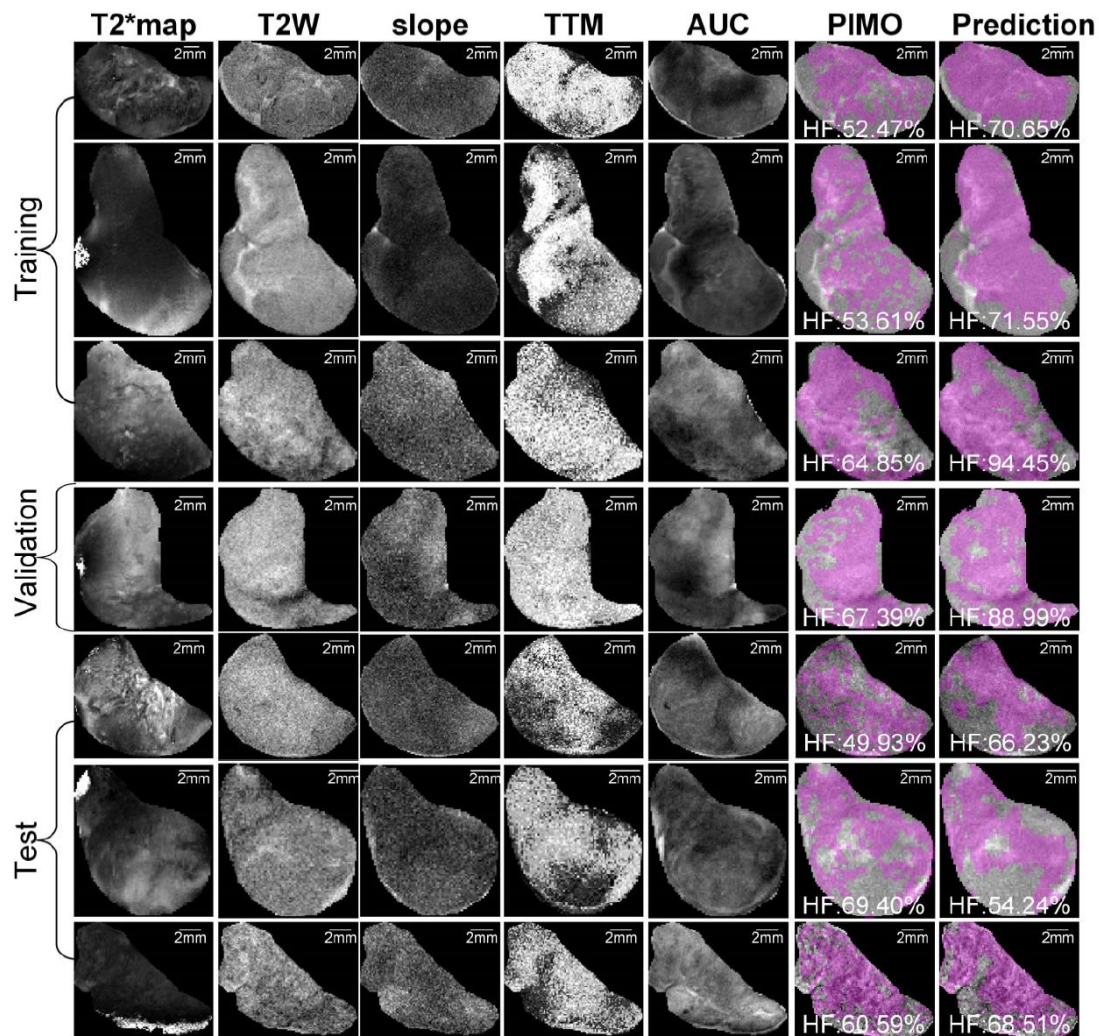


# AI trends

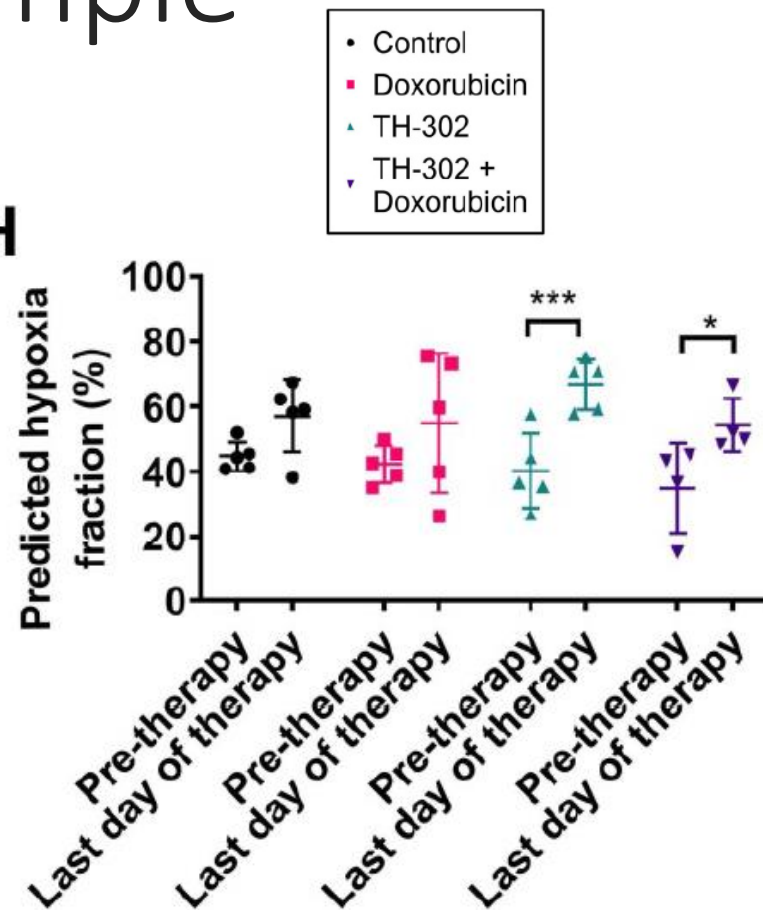
AI PubMed Searches by Year



# Imaging - example

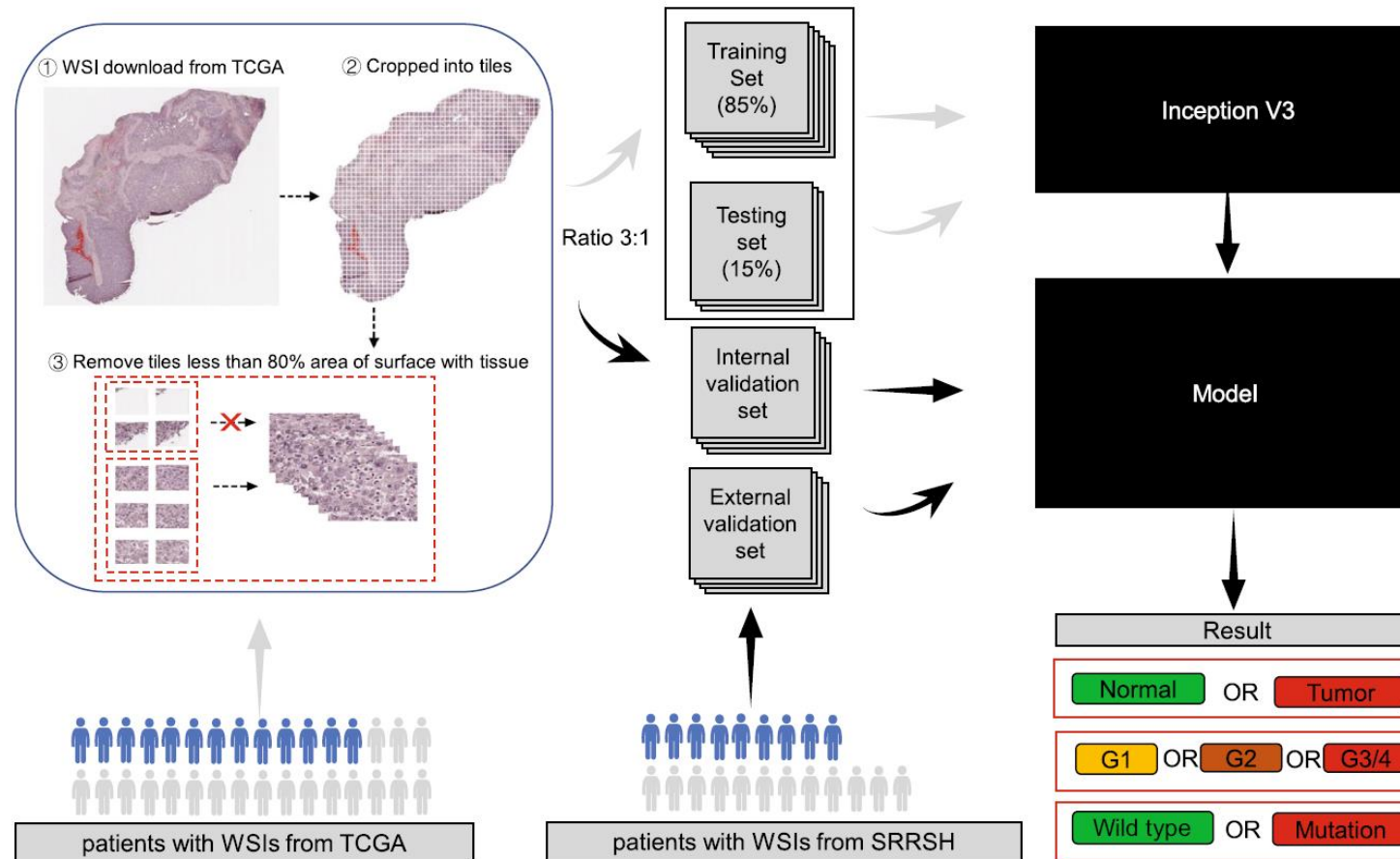


H



# Digital pathology - example

Hepatocellular carcinoma



“CTNNB1, FMN2, TP53, and ZFX4 mutations, could be predicted from histopathology images, with external AUCs from 0.71 to 0.89.”

Chen M, et al., *npj Precision Oncology*, 2020

# Insurance claims analysis – example

Poland 2009-2015, Cancer related contacts



3.7 million of patients



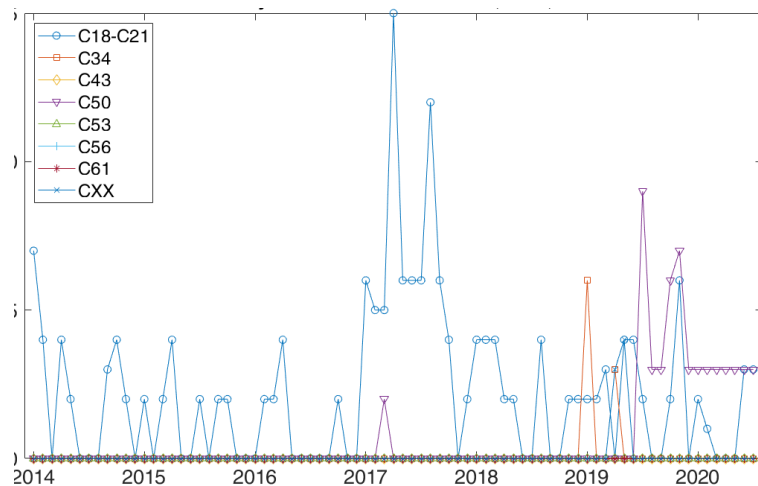
83.3 million of contacts with healthcare system



114.8 million of insurance claims

There are too many errors in insurance claims, so additional cancer registry is crucial. However, we used AI to try to filter the data.

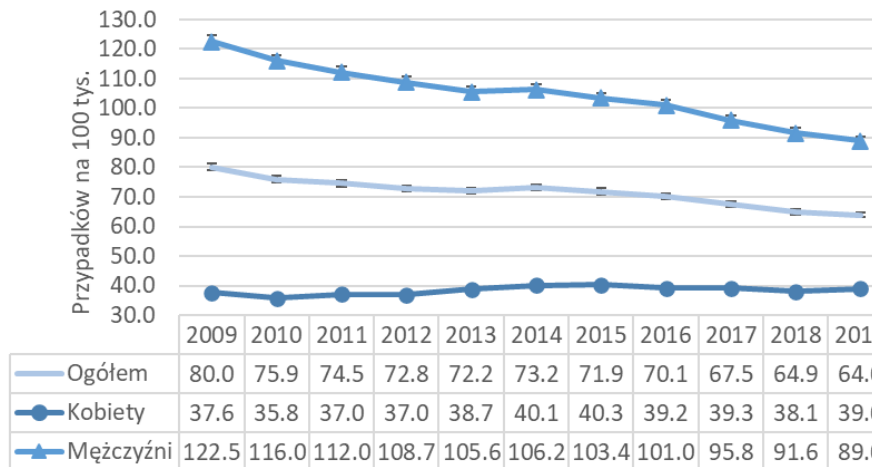
Patients' full history



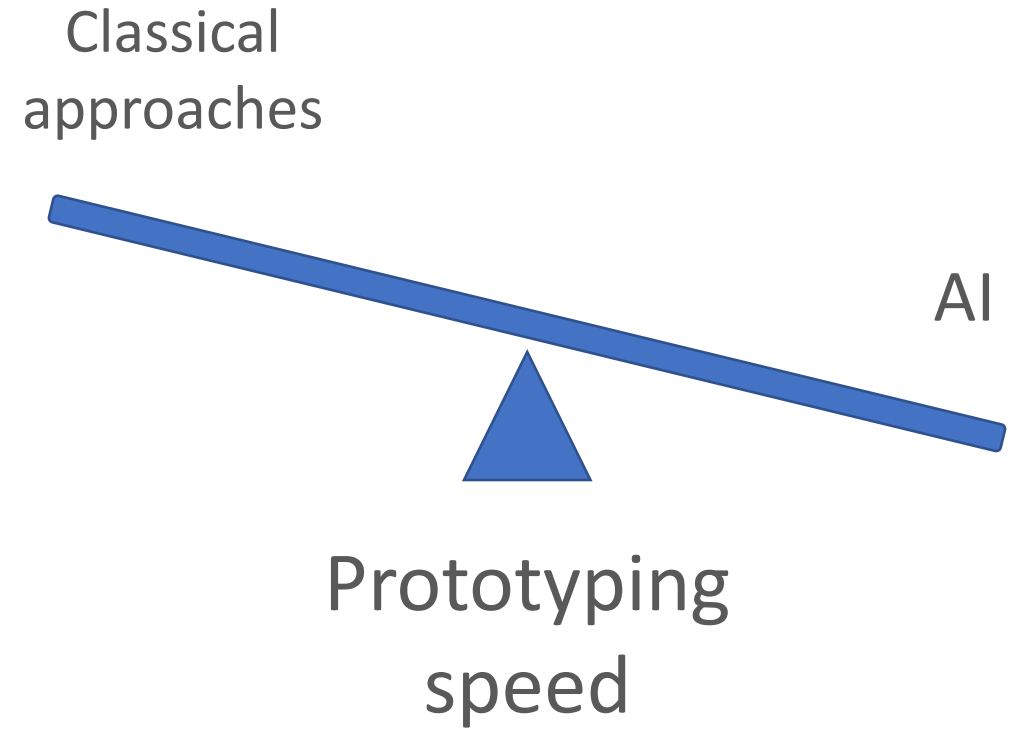
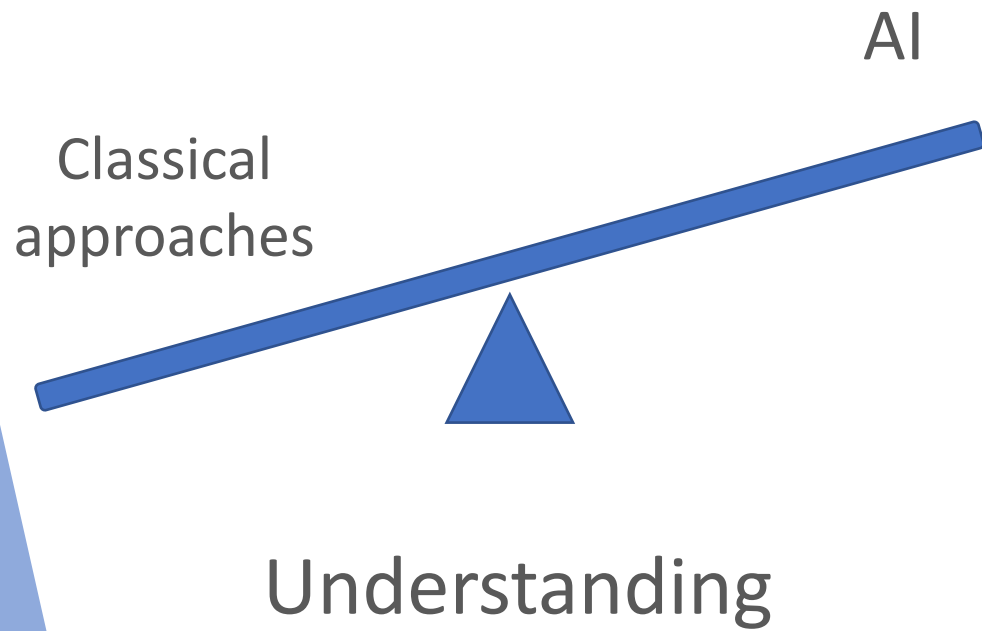
AI



Incidence

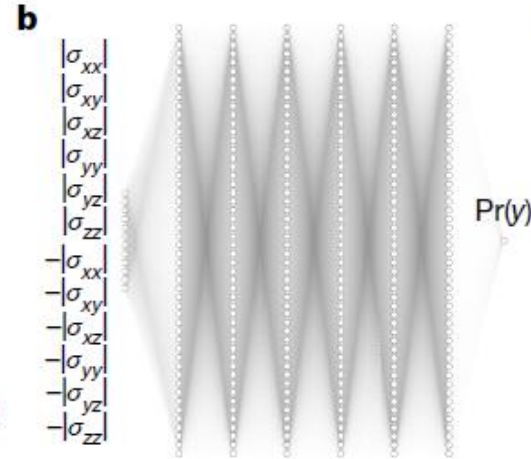
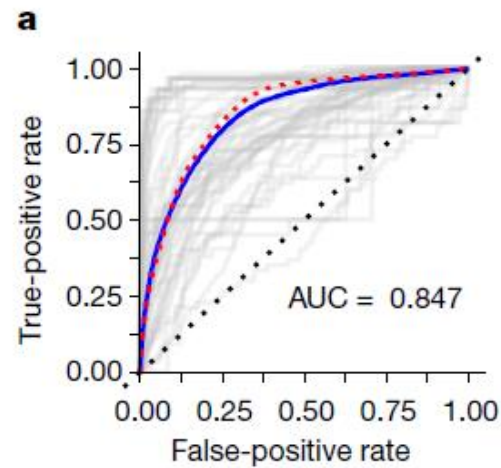
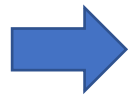


# There are some tradeoffs

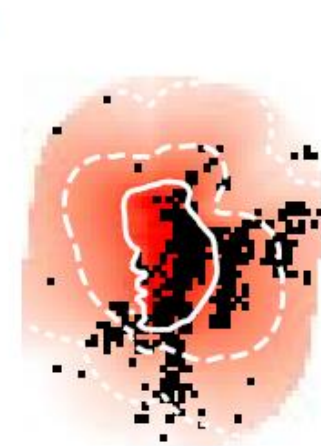
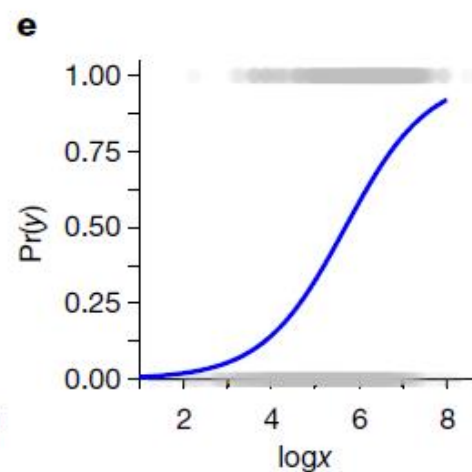
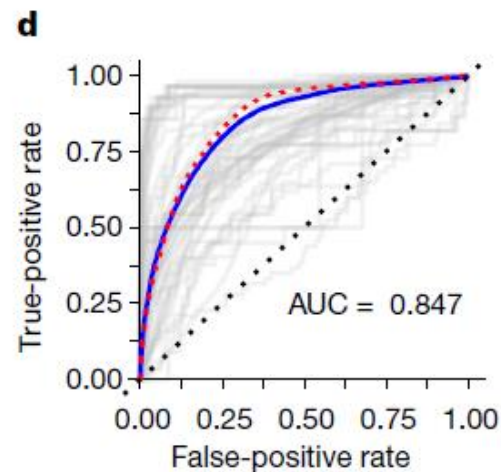
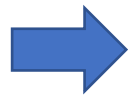


# Sometimes more doesn't mean better

13,451-parameter  
Deep Neural  
Network



2-parameter  
logistic  
regression



Mignan, A., Broccardo, M., *Nature* **574**, E1–E3  
(2019)

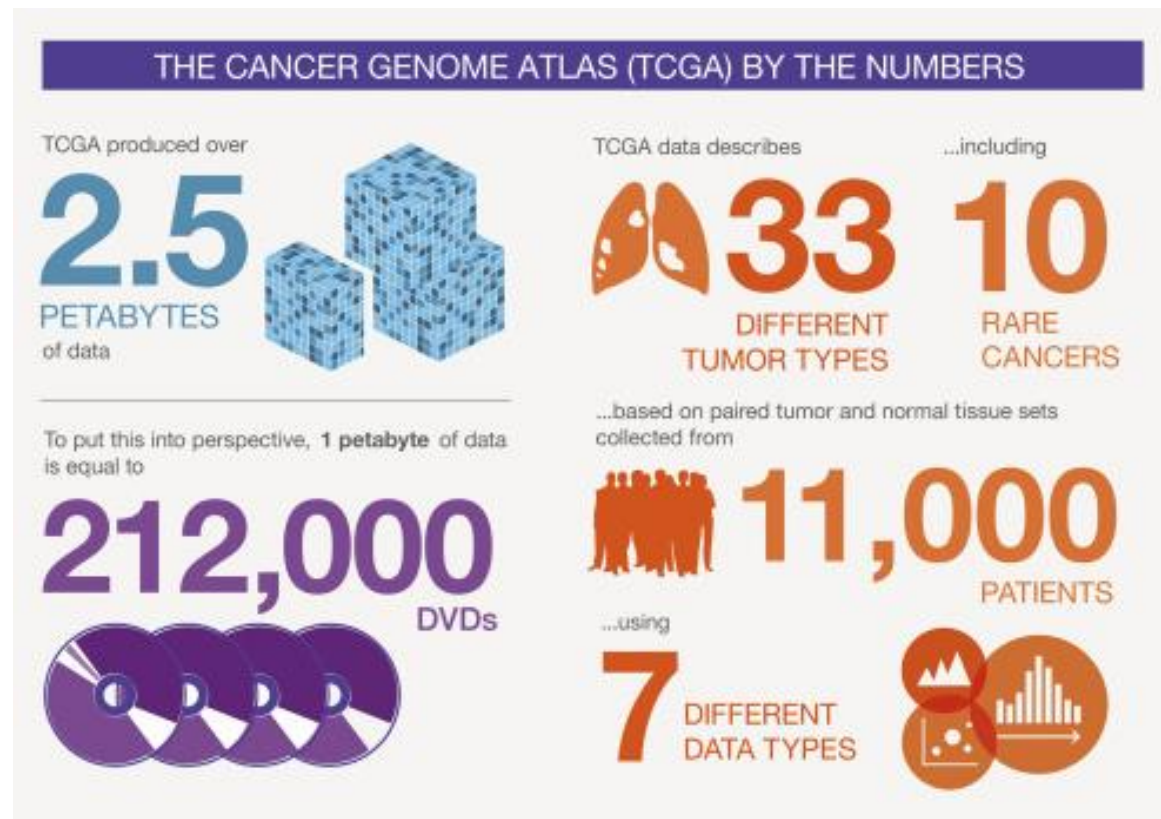
# Risk of bias when limited understanding

“Correctional Offender Management Profiling for Alternative Sanctions **biased against black people**”

“Amazon’s Recruiting Engine **biased against women**”

“Predictive Policing (PredPol) artificial intelligence **biased against minorities**”

# A lot of data per-patient, but not enough patients



Despite the huge amount of data in oncology, we face the problem of small cohorts for many cancer types.

AI algorithms frequently need many more patients than classical approaches (due to the sheer number of parameters).



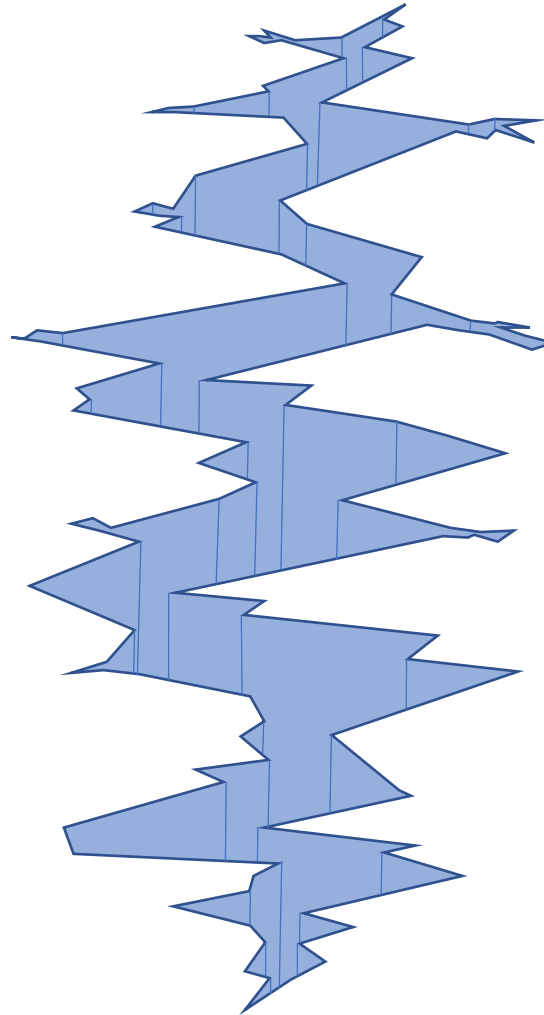
# Most important data in EHR, but...

## Electronic Health Records

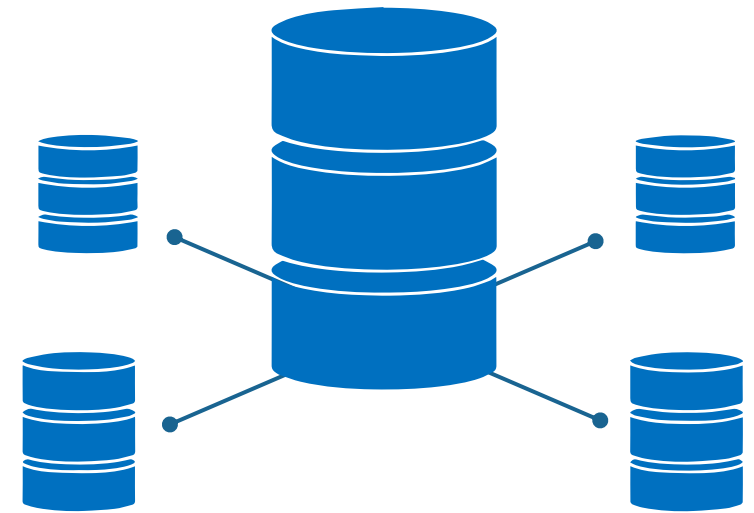


Focused on data required for reimbursement claims and regulatory bodies.

Essential clinical data mostly stored in unstructured format.



## Real World Data

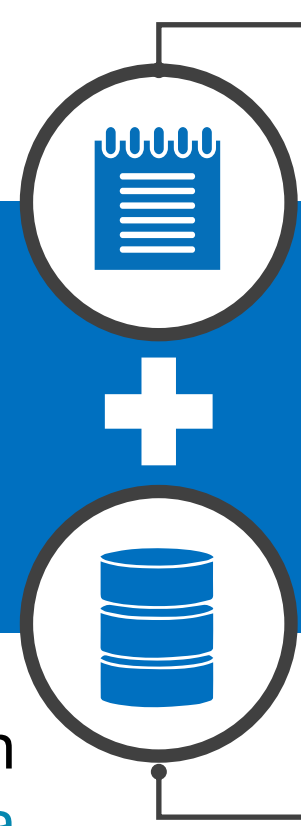


Key clinical data readily available for longitudinal analysis and interpretation (evidence).

# AI as a support tool to generate RWD – the future

## Live Suggestions For The Proper Content

Interaction with proprietary and continuously growing medical dictionaries through NLP algorithms



## Electronic Health Records

Complete documentation that is easily manageable

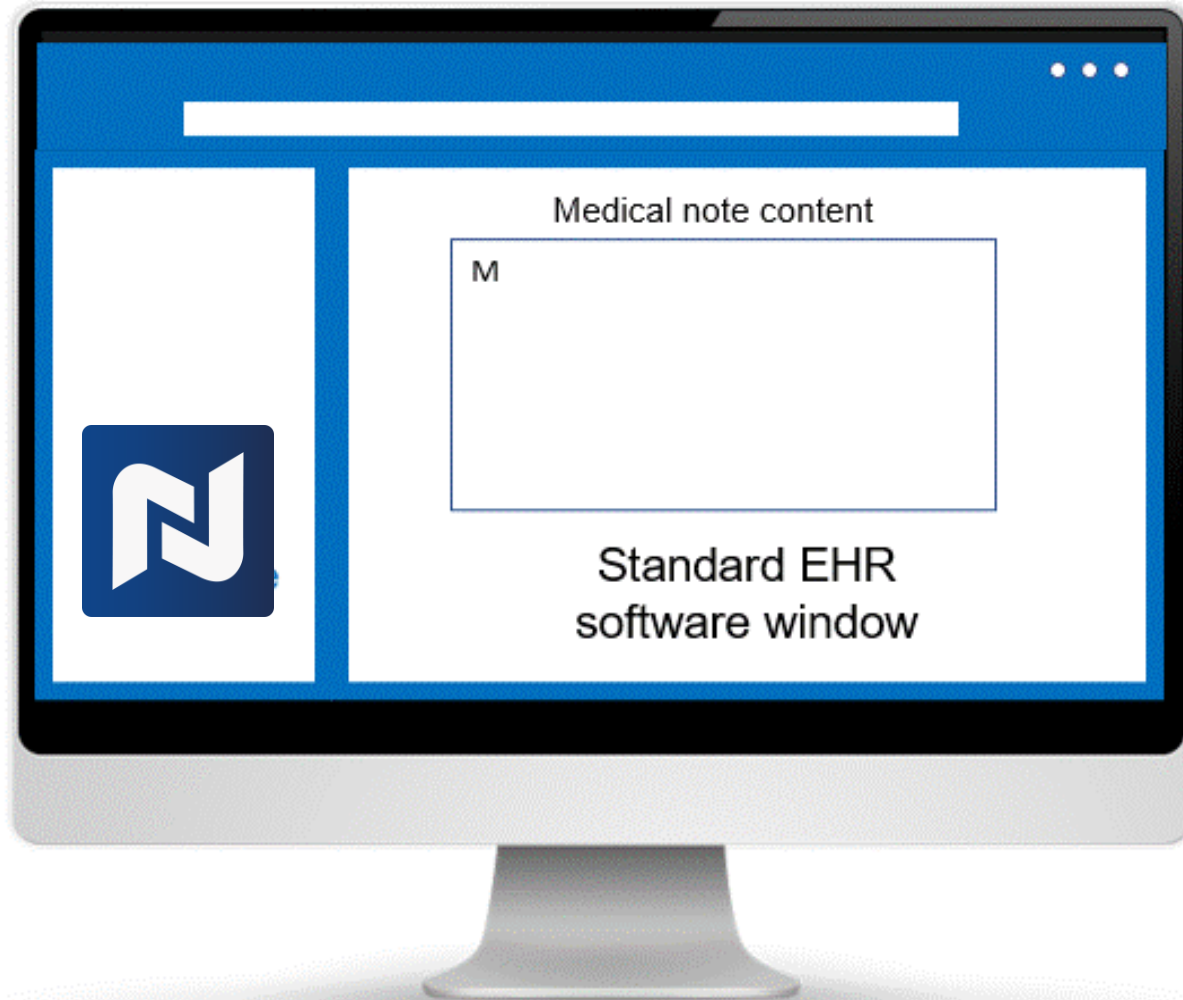


“Complete Electronic Health Records **And Real World Data** Created Simultaneously”

## Real World Data

Precise and complete database of clinical knowledge

# CliniNote technology in practice

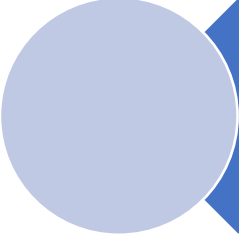


- ✓ Run your HIS as usual
- ✓ Use the hints from CN sidebar
- ✓ This is the D2D work of the doctors



“Currently in various hospitals in Poland”

# Summary



Classical statistical approaches won't be replaced but augmented by AI



Most AI approaches serve as a tool to generate additional information



In order to be successful we need to change the way in which health records are generated