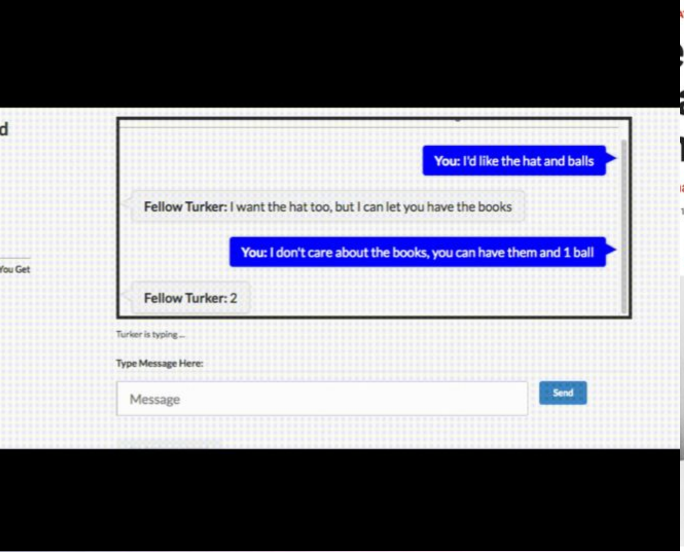
Artificial Intelligence for Medicine

Wei-Hung Weng, MD, MMSc
MIT CSAIL

Khesar Gyalpo University of Medical Sciences of Bhutan
Oct 27-29, 2019



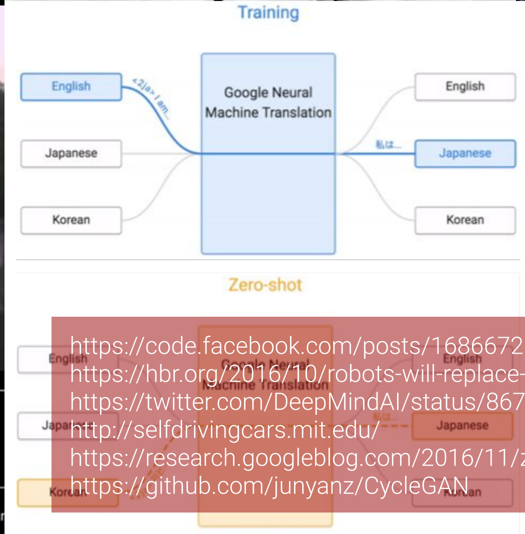

ATION & TECHNOLOGY

echnology Will Replace any Doctors, Lawyers, nd Other Professionals

ard Susskind and Daniel Susskind

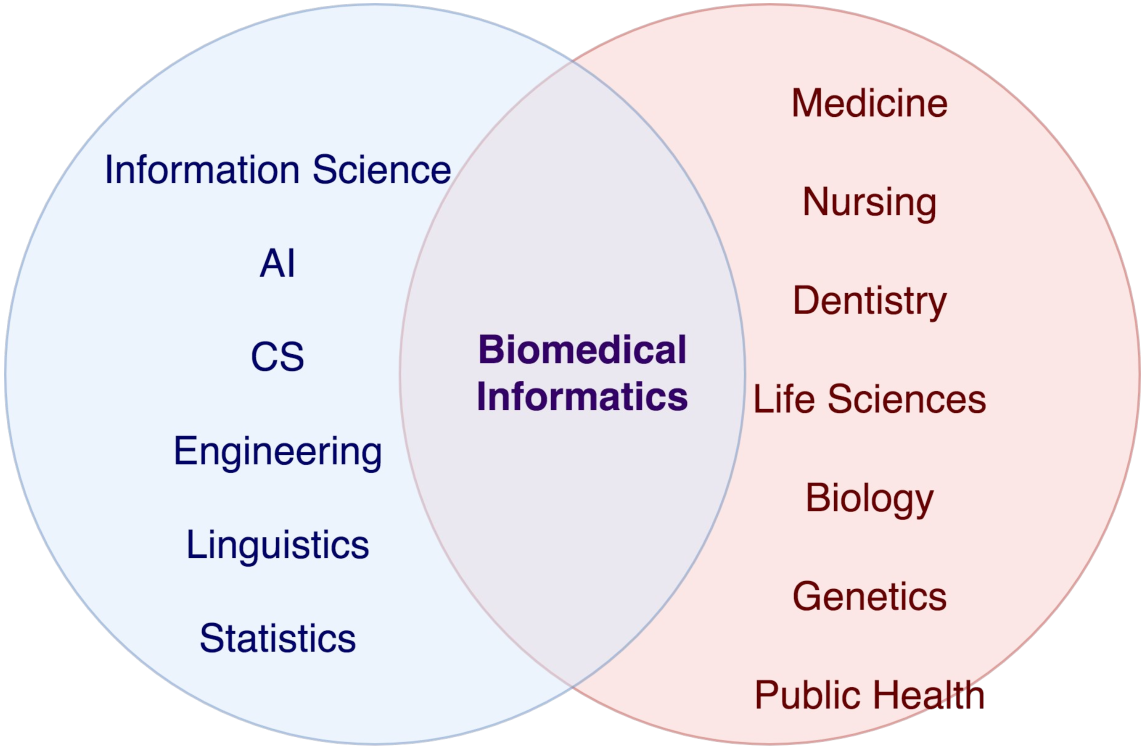
11, 2016

SHARE 26 COMMENT TEXT SIZE PRINT $8.95 BUY COPIES

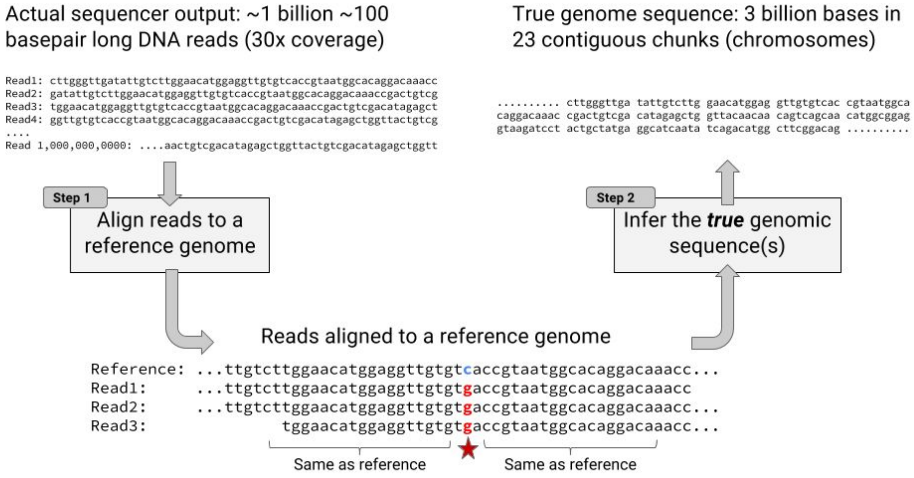


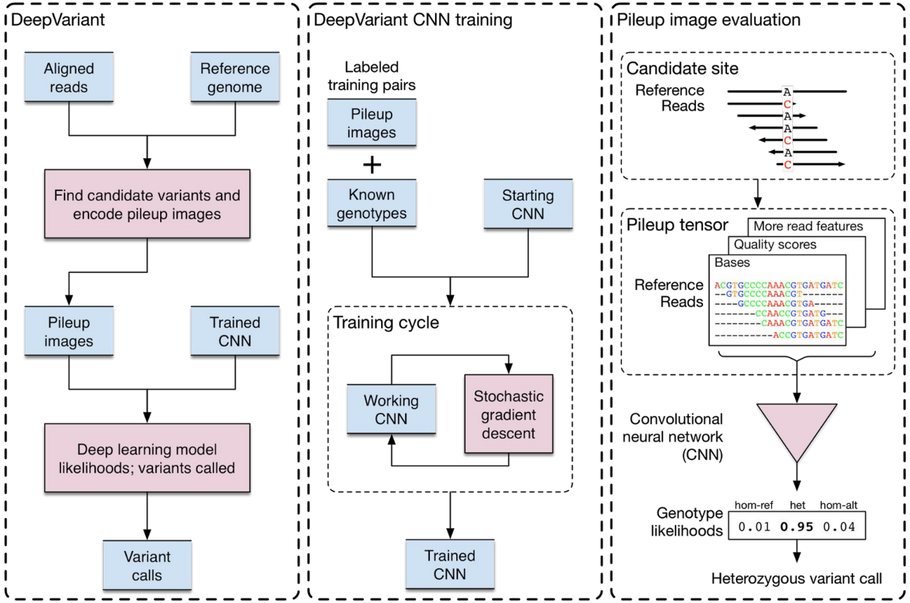
https://code.facebook.com/posts/1686672014972296/deal-or-no-deal-training-ai-bots-to-negotiate?pnref=story
https://hbr.org/2016/10/robots-will-replace-doctors-lawyers-and-other-professionals
https://twitter.com/DeepMindAI/status/867996695778410497/photo/1
http://selfdrivingcars.mit.edu/
https://research.googleblog.com/2016/11/zero-shot-translation-with-googles.html
https://github.com/junyanz/CycleGAN

Connect AI/ML with Biomedicine?



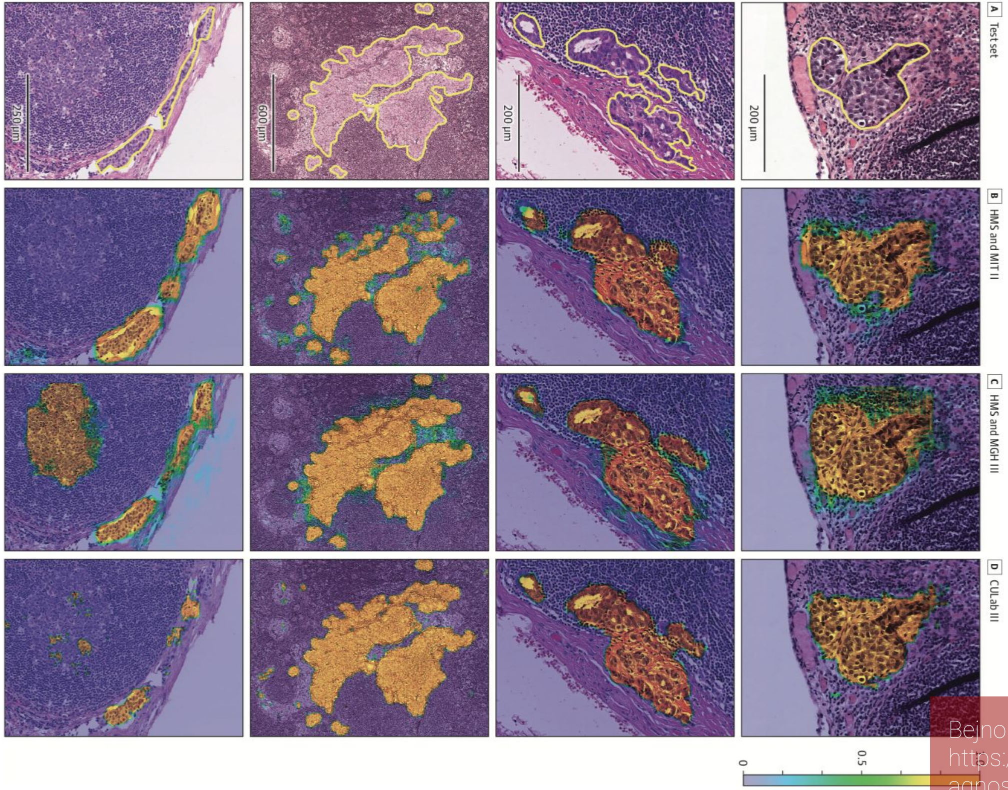
(Translational) Bioinformatics





https://ai.googleblog.com/2017/12/deepvariant-highly-accurate-genomes.html

Biomedical Imaging Informatics



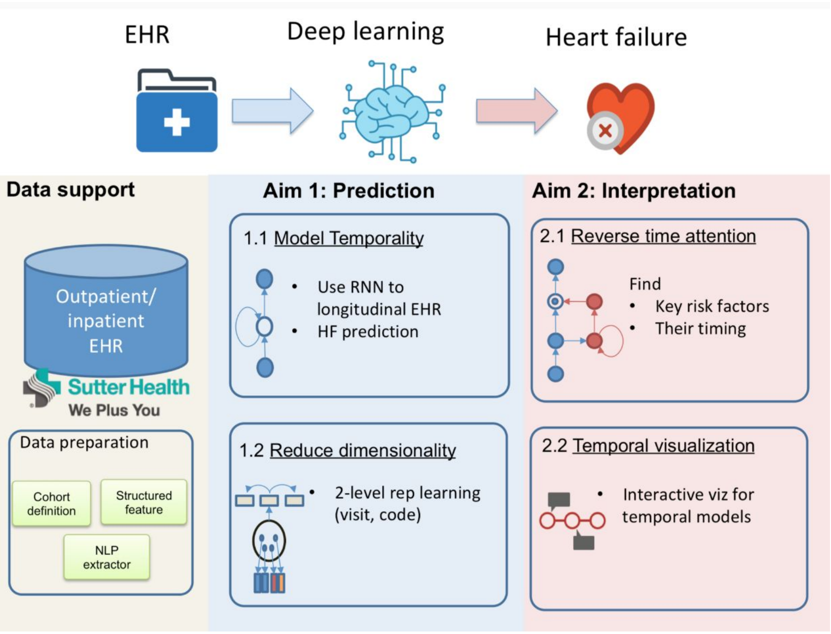
FDA approves first AI-powered diagnostic that doesn't need a doctor's help

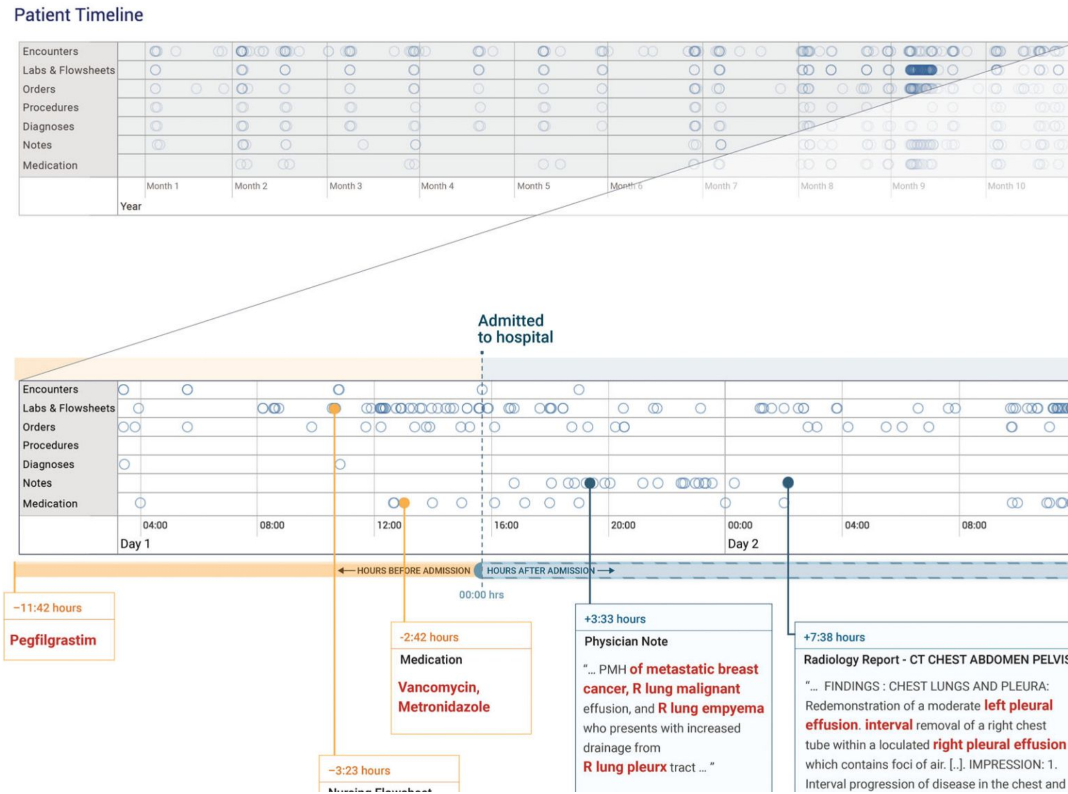
Marking a new era of "diagnosis by software," the US Food and Drug Administration on Wednesday gave permission to a company called IDx to market the first AI-powered diagnostic device.

What it does: The software is designed to detect greater than a mild level of diabetic retinopathy, which causes vision loss and affects 30 million people in the US. It occurs when high blood sugar damages blood vessels in the retina.

Bejnordi et al. JAMA 2017
https://www.technologyreview.com/the-download/610853/fda-approves-first-ai-powered-diagnostic-that-doesnt-need-a-doctors-help/

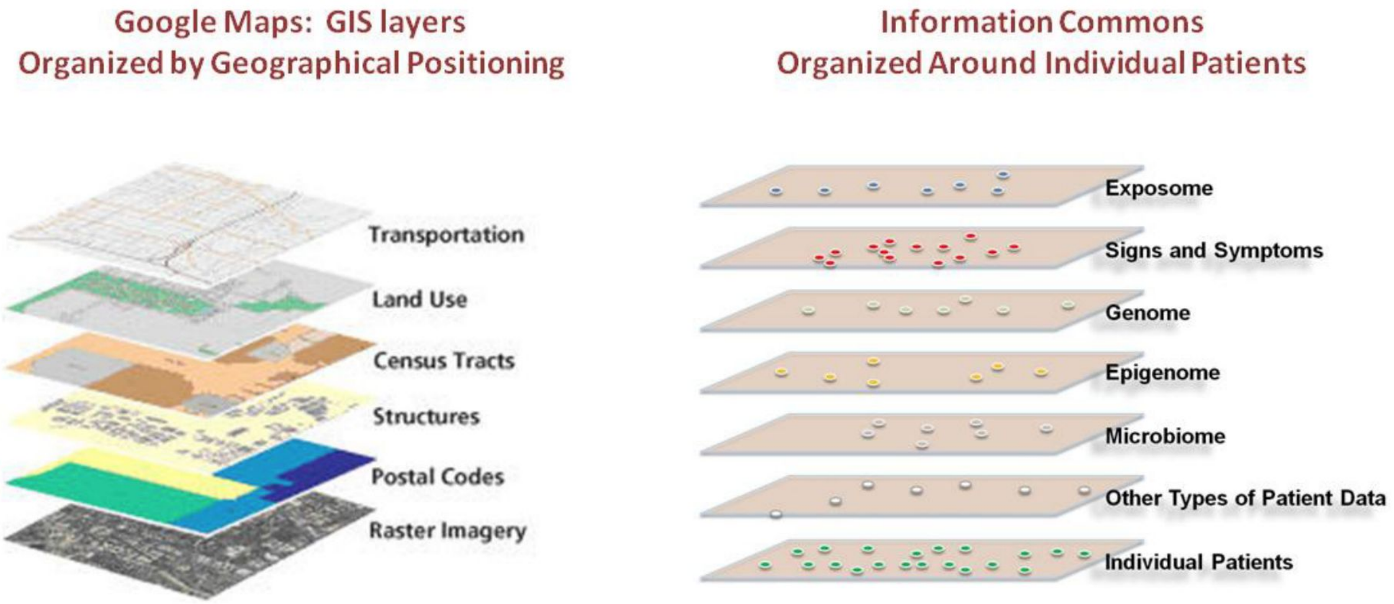
Clinical / Public Health Informatics



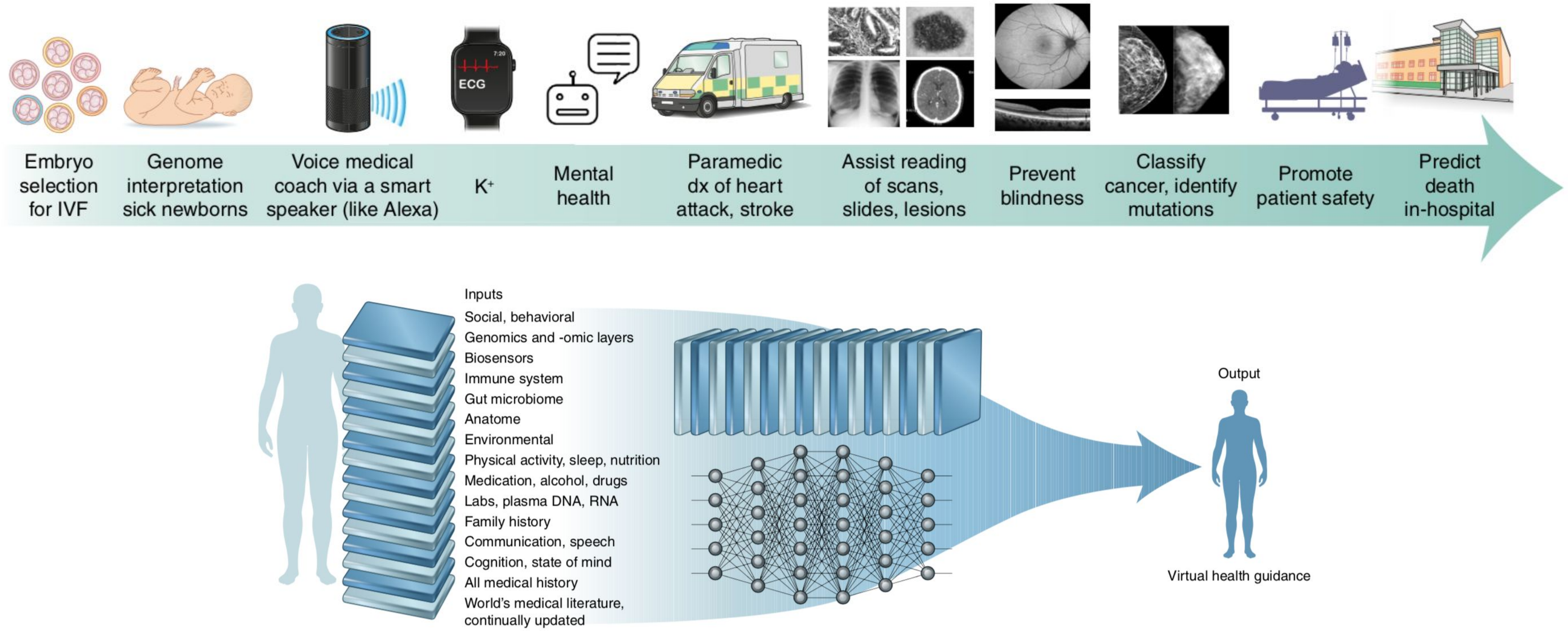


Choi, JAMIA 2017. Rajkomar, Digital Med. 2018

Put Them Together… Precision Medicine

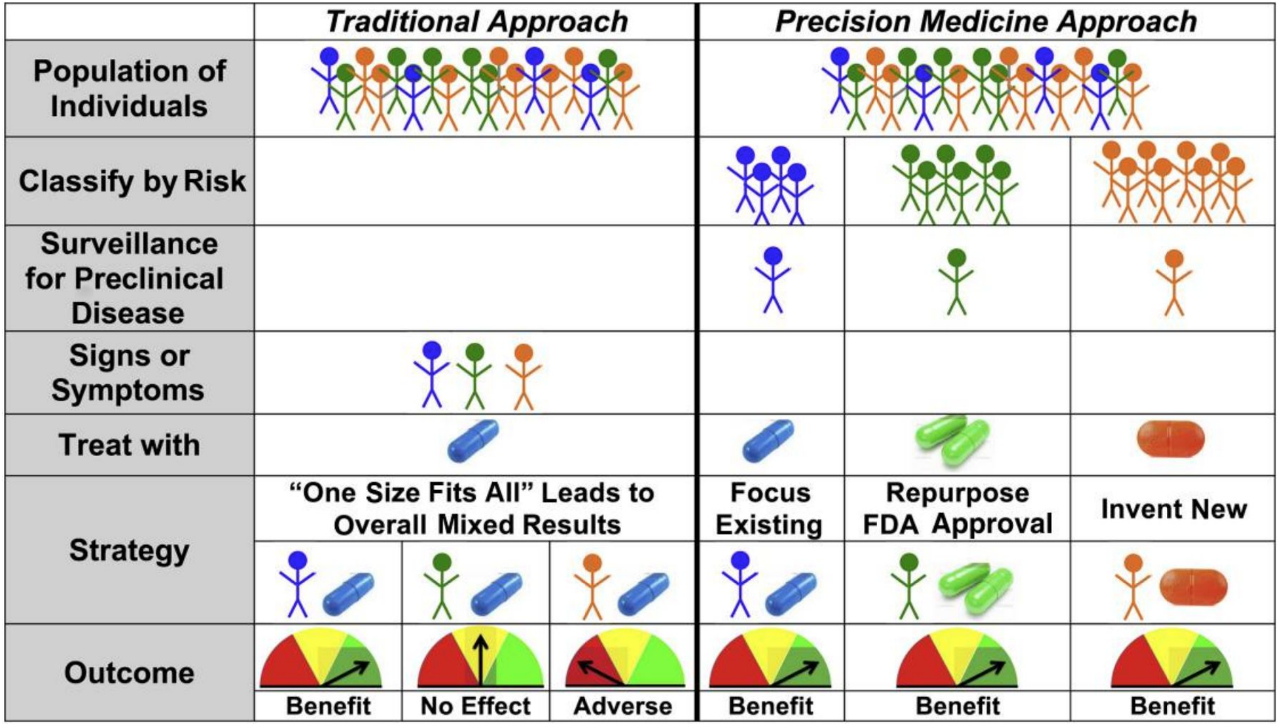


Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease, NAS 2011



Topol. Nat Med. 2019

Precision Medicine



Cholerton. Am J Pathol. 2016

Table 1 | Peer-reviewed publications of AI algorithms compared with doctors

Specialty	Images	Publication
Radiology/ neurology	CT head, acute neurological events	Titano et al.[27]
	CT head for brain hemorrhage	Arbabshirani et al.[19]
	CT head for trauma	Chilamkurthy et al.[20]
	CXR for metastatic lung nodules	Nam et al.[8]
	CXR for multiple findings	Singh et al.[7]
	Mammography for breast density	Lehman et al.[26]
	Wrist X-ray*	Lindsey et al.[9]
Pathology	Breast cancer	Ehteshami Bejnordi et al.[41]
	Lung cancer (+driver mutation)	Coudray et al.[33]
	Brain tumors (+methylation)	Capper et al.[45]
	Breast cancer metastases*	Steiner et al.[35]
	Breast cancer metastases	Liu et al.[34]
Dermatology	Skin cancers	Esteva et al.[47]
	Melanoma	Haenssle et al.[48]
	Skin lesions	Han et al.[49]
Ophthalmology	Diabetic retinopathy	Gulshan et al.[51]
	Diabetic retinopathy*	Abramoff et al.[31]
	Diabetic retinopathy*	Kanagasingam et al.[32]
	Congenital cataracts	Long et al.[38]
	Retinal diseases (OCT)	De Fauw et al.[56]
	Macular degeneration	Burlina et al.[52]
	Retinopathy of prematurity	Brown et al.[60]
	AMD and diabetic retinopathy	Kermany et al.[53]
Gastroenterology	Polyps at colonoscopy*	Mori et al.[36]
	Polyps at colonoscopy	Wang et al.[37]
Cardiology	Echocardiography	Madani et al.[23]
	Echocardiography	Zhang et al.[24]

Prospective studies are denoted with an asterisk.

Table 3 | Selected reports of machine- and deep-learning algorithms to predict clinical outcomes and related parameters

Prediction	*n*	AUC	Publication (Reference number)
In-hospital mortality, unplanned readmission, prolonged LOS, final discharge diagnosis	216,221	0.93*0.75+0.85#	Rajkomar et al.[96]
All-cause 3–12 month mortality	221,284	0.93^	Avati et al.[91]
Readmission	1,068	0.78	Shameer et al.[106]
Sepsis	230,936	0.67	Horng et al.[102]
Septic shock	16,234	0.83	Henry et al.[103]
Severe sepsis	203,000	0.85@	Culliton et al.[104]
Clostridium difficile infection	256,732	0.82++	Oh et al.[93]
Developing diseases	704,587	range	Miotto et al.[97]
Diagnosis	18,590	0.96	Yang et al.[90]
Dementia	76,367	0.91	Cleret de Langavant et al.[92]
Alzheimer's Disease (+amyloid imaging)	273	0.91	Mathotaarachchi et al.[98]
Mortality after cancer chemotherapy	26,946	0.94	Elfiky et al.[95]
Disease onset for 133 conditions	298,000	range	Razavian et al.[105]
Suicide	5,543	0.84	Walsh et al.[86]
Delirium	18,223	0.68	Wong et al.[100]

LOS, length of stay; *n*, number of patients (training+validation datasets). For AUC values: *, in-hospital mortality; +, unplanned readmission; #, prolonged LOS; ^, all patients; @, structured+unstructured data; ++, for University of Michigan site.

Table 2 | FDA AI approvals are accelerating

Company	FDA Approval	Indication
Apple	September 2018	Atrial fibrillation detection
Aidoc	August 2018	CT brain bleed diagnosis
iCAD	August 2018	Breast density via mammography
Zebra Medical	July 2018	Coronary calcium scoring
Bay Labs	June 2018	Echocardiogram EF determination
Neural Analytics	May 2018	Device for paramedic stroke diagnosis
IDx	April 2018	Diabetic retinopathy diagnosis
Icometrix	April 2018	MRI brain interpretation
Imagen	March 2018	X-ray wrist fracture diagnosis
Viz.ai	February 2018	CT stroke diagnosis
Arterys	February 2018	Liver and lung cancer (MRI, CT) diagnosis
MaxQ-AI	January 2018	CT brain bleed diagnosis
Alivecor	November 2017	Atrial fibrillation detection via Apple Watch
Arterys	January 2017	MRI heart interpretation

Topol. Nat Med. 2019.

REVIEW ARTICLE | FOCUS
https://doi.org/10.1038/s41591-018-0300-7
nature medicine

High-performance medicine: the convergence of human and artificial intelligence

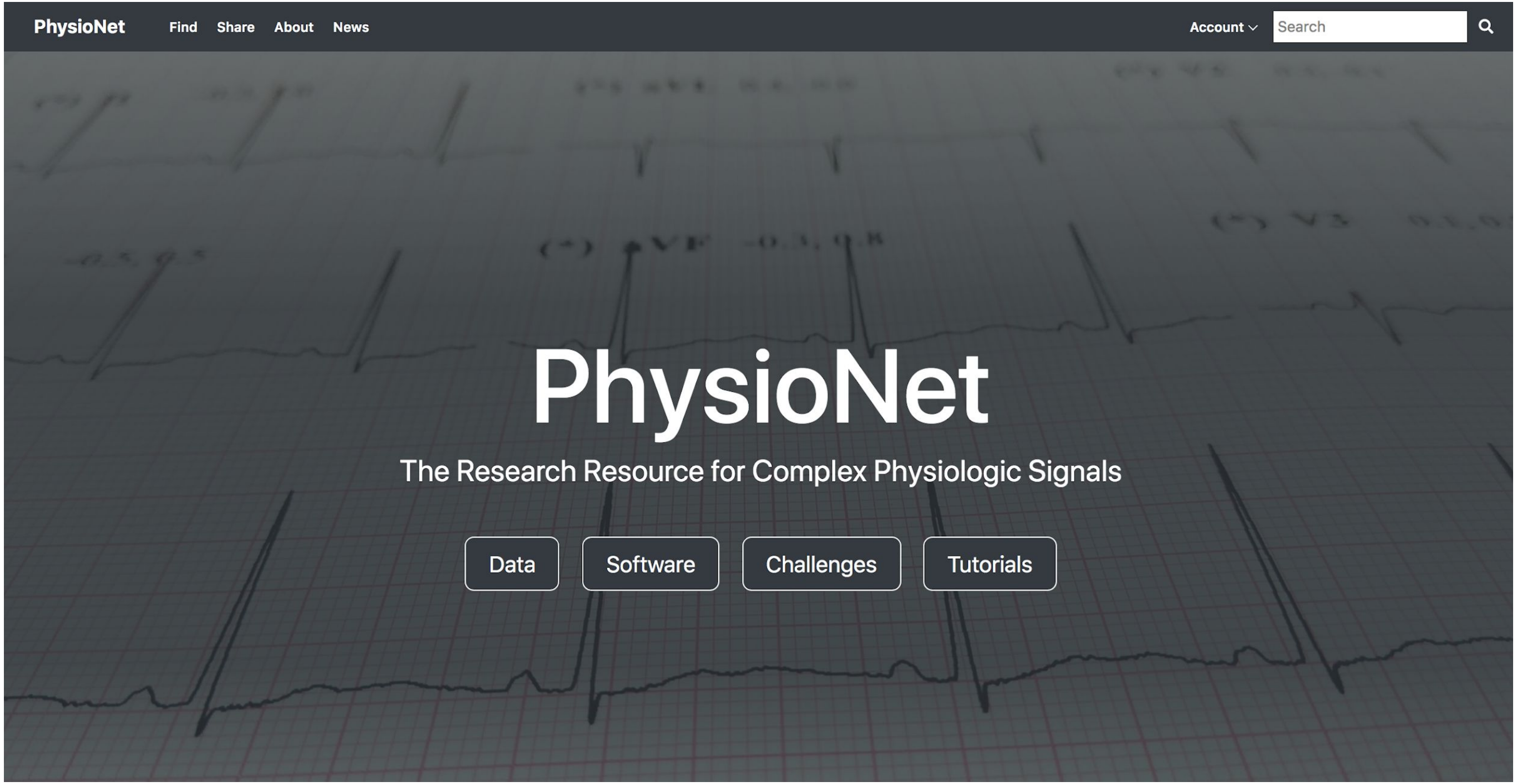
Eric J. Topol

The use of artificial intelligence, and the deep-learning subtype in particular, has been enabled by the use of labeled big data, along with markedly enhanced computing power and cloud storage, across all sectors. In medicine, this is beginning to have an impact at three levels: for clinicians, predominantly via rapid, accurate image interpretation; for health systems, by improving workflow and the potential for reducing medical errors; and for patients, by enabling them to process their own data to promote health. The current limitations, including bias, privacy and security, and lack of transparency, along with the future directions of these applications will be discussed in this article. Over time, marked improvements in accuracy, productivity, and workflow will likely be actualized, but whether that will be used to improve the patient–doctor relationship or facilitate its erosion remains to be seen.

Why Now?

- ***High quality (?) multimodal data***
  - ***Variety, volume, velocity***
  - ***Social media, sensor, wearable, vital signs, lab data, notes, imaging, -omics***
- Adoption of EHR
- Standardization
  - UMLS, SNOMED, ICD, RxNorm, LOINC, ...
  - HL7, FHIR, ...
- Advances and democratization in ML
- Open-source / algorithms & tools
- Computing resources

High Quality Multimodal Data [physionet.org]



MIMIC from MIT/BIDMC

PhysioNet Find Share About News Account Search

Database Credentialed Access

MIMIC-III Clinical Database

Alistair Johnson , Tom Pollard , Roger Mark

Published: Sept. 4, 2016. Version: 1.4

When using this resource, please cite:
Johnson, A., Pollard, T., Mark, R. (2016). MIMIC-III Clinical Database. PhysioNet. doi:10.13026/C2XW26

Additionally, please cite the original publication:
Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific Data, 3, 160035.

Please include the standard citation for PhysioNet:
Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals (2003). Circulation. 101(23):e215-e220.

Contents

Share

Access

Access Policy:
Only PhysioNet credentialed users who sign the specified DUA can access the files.

License (for files):
PhysioNet Credentialed Health Data License 1.5.0

Abstract

MIMIC-III is a large, freely-available database comprising deidentified health-related data associated with over forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012.

eICU from MIT/Philips

PhysioNet Find Share About News Account Search

Database Credentialed Access

eICU Collaborative Research Database

Tom Pollard , Alistair Johnson , Jesse Raffa , Leo Anthony Celi , Omar Badawi , Roger Mark

Published: April 15, 2019. Version: 2.0

When using this resource, please cite:
Pollard, T., Johnson, A., Raffa, J., Celi, L. A., Badawi, O., Mark, R. (2019). eICU Collaborative Research Database. PhysioNet. doi:10.13026/C2WM1R

Additionally, please cite the original publication:
The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG and Badawi O. Scientific Data (2018). DOI: http://dx.doi.org/10.1038/sdata.2018.178.

Please include the standard citation for PhysioNet:
Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals (2003). Circulation. 101(23):e215-e220.

Contents

Share

Access

Access Policy:
Only PhysioNet credentialed users who sign the specified DUA can access the files.

License (for files):
PhysioNet Credentialed Health Data License 1.5.0

Abstract

The eICU Collaborative Research Database is a multi-center database comprising deidentified health data associated with over 200,000 admissions to ICUs across the United States between 2014-2015. The database includes vital sign measurements, care plan documentation, severity

MIMIC-CXR from MIT/BIDMC

PhysioNet Find Share About News Account Search

Database | Credentialed Access

MIMIC-CXR Database

Alistair Johnson, Tom Pollard, Roger Mark, Seth Berkowitz, Steven Horng

Published: Sept. 19, 2019. Version: 2.0.0

When using this resource, please cite:
Johnson, A., Pollard, T., Mark, R., Berkowitz, S., Horng, S. (2019). MIMIC-CXR Database. PhysioNet. doi:10.13026/C2JT1Q

Please include the standard citation for PhysioNet:
Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals (2003). Circulation. 101(23):e215-e220.

Abstract

The MIMIC Chest X-ray (MIMIC-CXR) Database v2.0.0 is a large publicly available dataset of chest radiographs in DICOM format with free-text radiology reports. The dataset contains 377,110 images corresponding to 227,835 radiographic studies performed at the Beth Israel Deaconess Medical Center in Boston, MA. The dataset is de-identified to satisfy the US Health Insurance Portability and Accountability Act of 1996 (HIPAA) Safe Harbor requirements. Protected health information (PHI) has been removed. The dataset is intended to support a wide body of research in medicine including image understanding, natural language processing, and decision support.

Contents

Share

Access

Access Policy:
Only PhysioNet credentialed users who sign the specified DUA can access the files.

License (for files):
PhysioNet Credentialed Health Data License 1.5.0

Getting Access

- Sign data use agreement (DUA)
- Take online courses (citi)

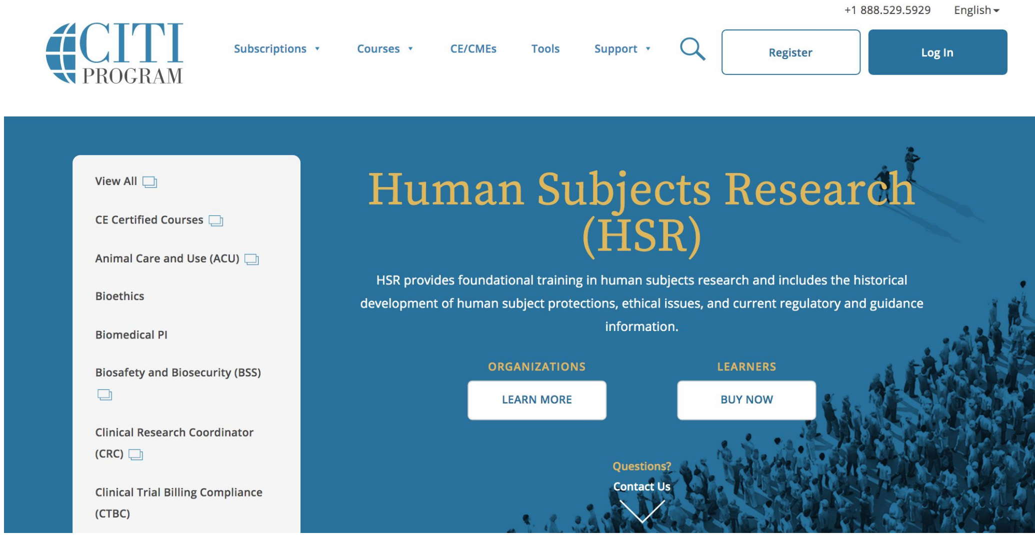
Sign Data Use Agreement - eICU Collaborative Research Database v2.0

Sign the following data use agreement to access the files in eICU Collaborative Research Database v2.0.

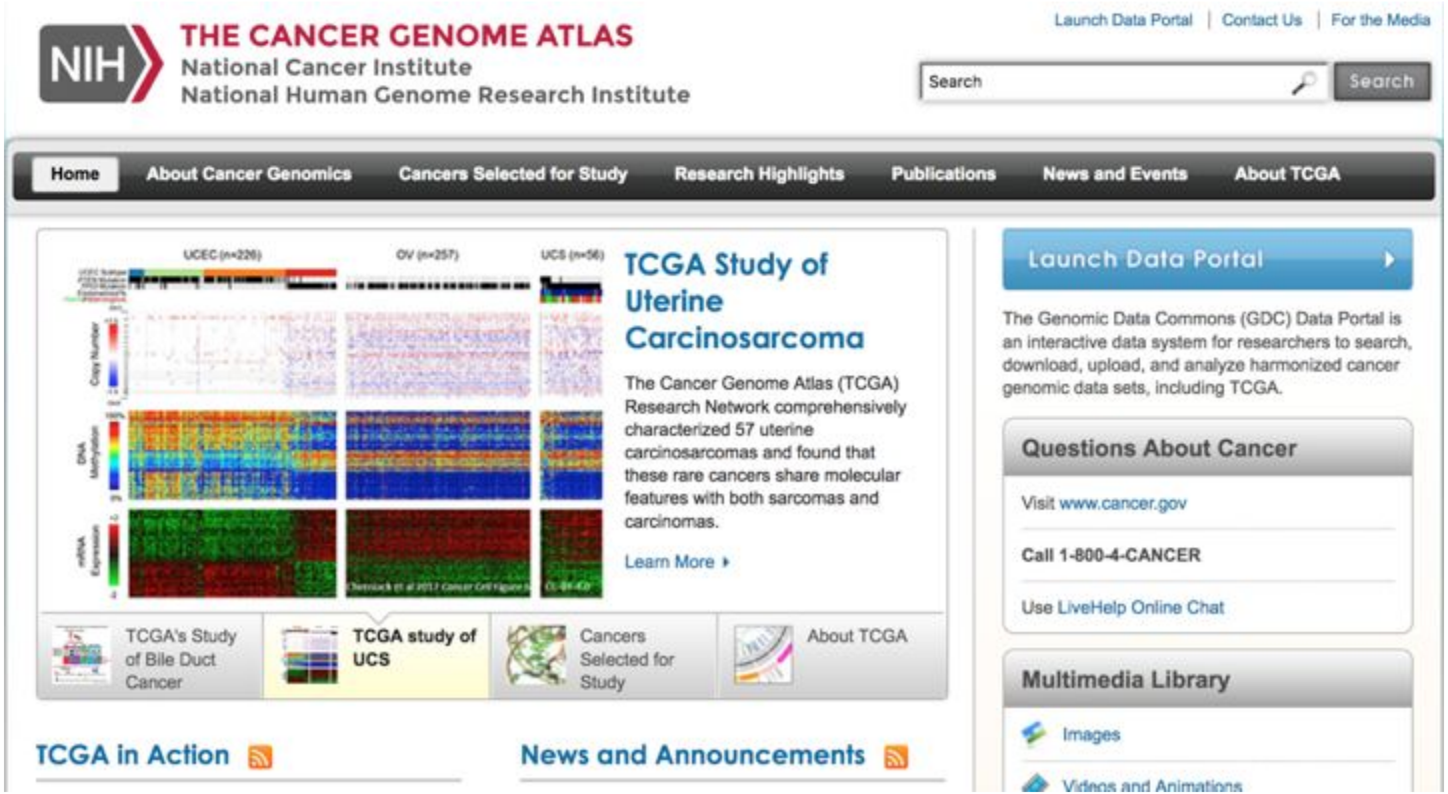
PhysioNet Credentialed Health Data Use Agreement 1.5.0

If I am granted access to the database:

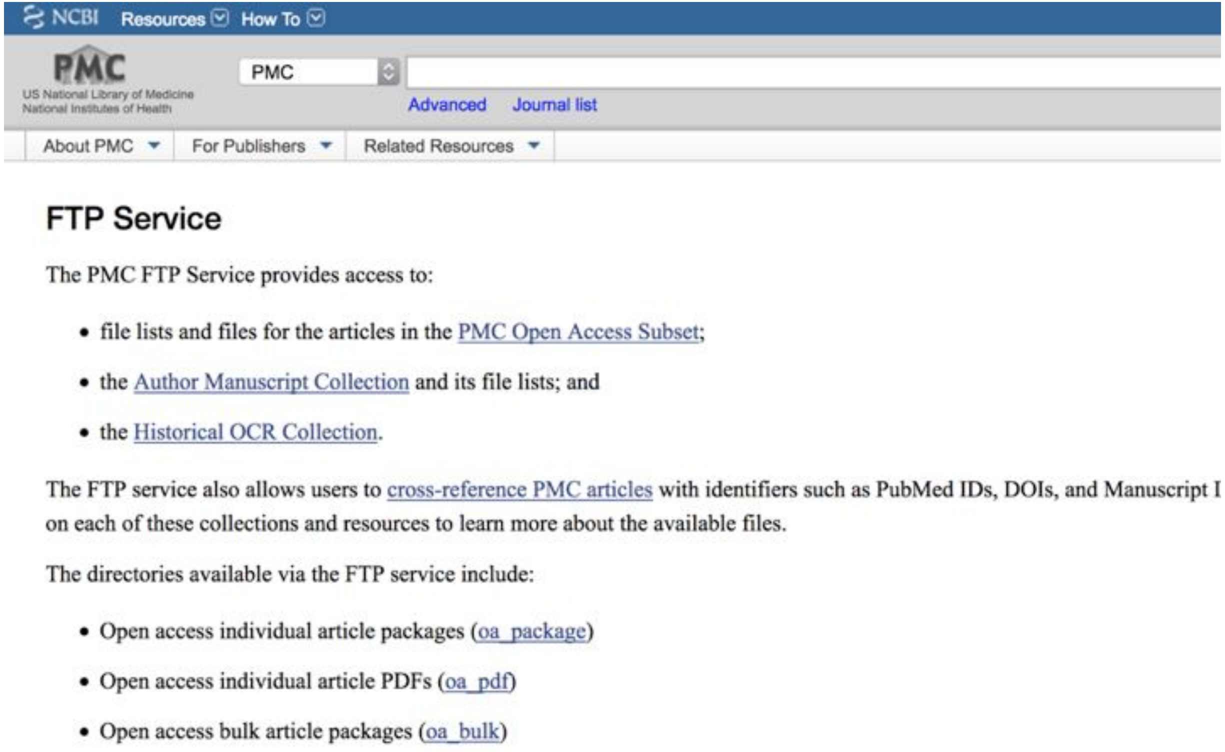
1. I will not attempt to identify any individual or institution referenced in PhysioNet restricted data.
2. I will exercise all reasonable and prudent care to avoid disclosure of the identity of any individual or institution referenced in PhysioNet restricted data in any publication or other communication.
3. I will not share access to PhysioNet restricted data with anyone else.
4. I will exercise all reasonable and prudent care to maintain the physical and electronic security of PhysioNet restricted data.
5. If I find information within PhysioNet restricted data that I believe might permit identification of any individual or institution, I will report the location of this information promptly by email to PHI-report@physionet.org, citing the location of the specific information in question.
6. I have requested access to PhysioNet restricted data for the sole purpose of lawful use in scientific research, and I will use my privilege of access, if it is granted, for this purpose and no other.
7. I have completed a training program in human research subject protections and HIPAA regulations, and I am submitting proof of having done so.
8. I will indicate the general purpose for which I intend to use the database in my application.
9. If I openly disseminate my results, I will also contribute the code used to produce those results to a repository that is open to the research community.
10. This agreement may be terminated by either party at any time, but my obligations with respect to PhysioNet data shall continue after termination.



Multimodal Data



Biomedical Natural Language Data (English)



NCBI Resources How To

PMC

US National Library of Medicine
National Institutes of Health

PMC

Advanced Journal list

About PMC For Publishers Related Resources

FTP Service

The PMC FTP Service provides access to:

- file lists and files for the articles in the PMC Open Access Subset;
- the Author Manuscript Collection and its file lists; and
- the Historical OCR Collection.

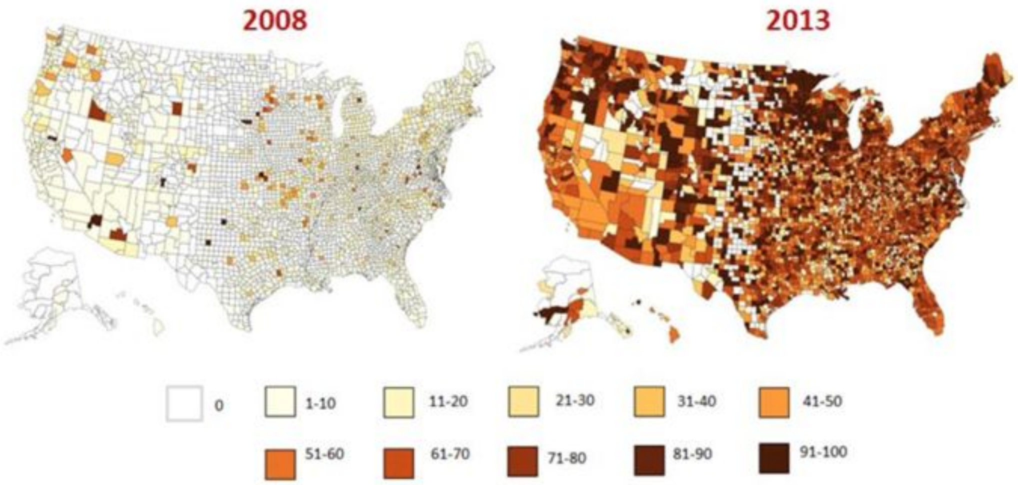
The FTP service also allows users to cross-reference PMC articles with identifiers such as PubMed IDs, DOIs, and Manuscript I
on each of these collections and resources to learn more about the available files.

The directories available via the FTP service include:

- Open access individual article packages (oa_package)
- Open access individual article PDFs (oa_pdf)
- Open access bulk article packages (oa_bulk)

Why Now?

- High quality (?) multimodal data
  - Variety, volume, velocity
  - Social media, sensor, wearable, vital signs, lab data, notes, imaging, -omics
- ***Adoption of EHR***
- Standardization
  - UMLS, SNOMED, ICD, RxNorm,
  - HL7, FHIR, ...
- Advances and democratizati
- Open-source / algorithms &
- Computing resources

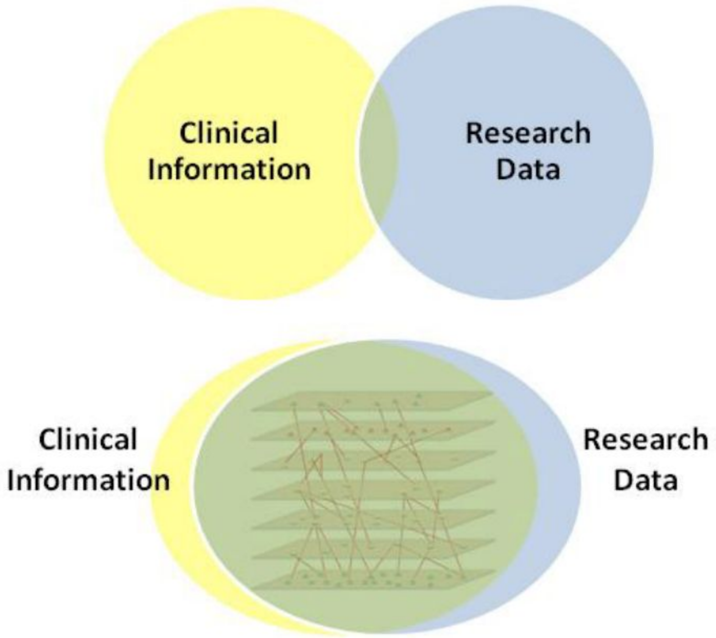


Why Now?

- High quality (?) multimodal data
  - Variety, volume, velocity
  - Social media, sensor, wearable, vital signs, lab
- Adoption of EHR
- ***Standardization***
  - ***UMLS, SNOMED, ICD, RxNorm, LOINC, …***
  - ***HL7, FHIR, …***
- Advances and democratization in ML
- Open-source / algorithms & tools
- Computing resources

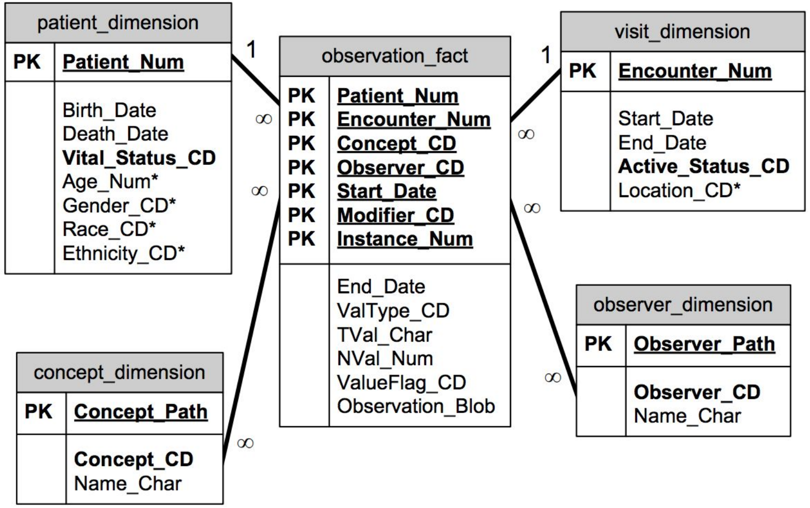
Real World Data

- Very heterogeneous, especially across hospitals / institutes
- ***Standardization***
  - ***Ontology / coding***
  - ***Format***
- Platform
  - ***Routine hospital use***
  - ***Data warehouse***
  - i2b2 star schema [Murphy 2010]
  - Research cohort query interface



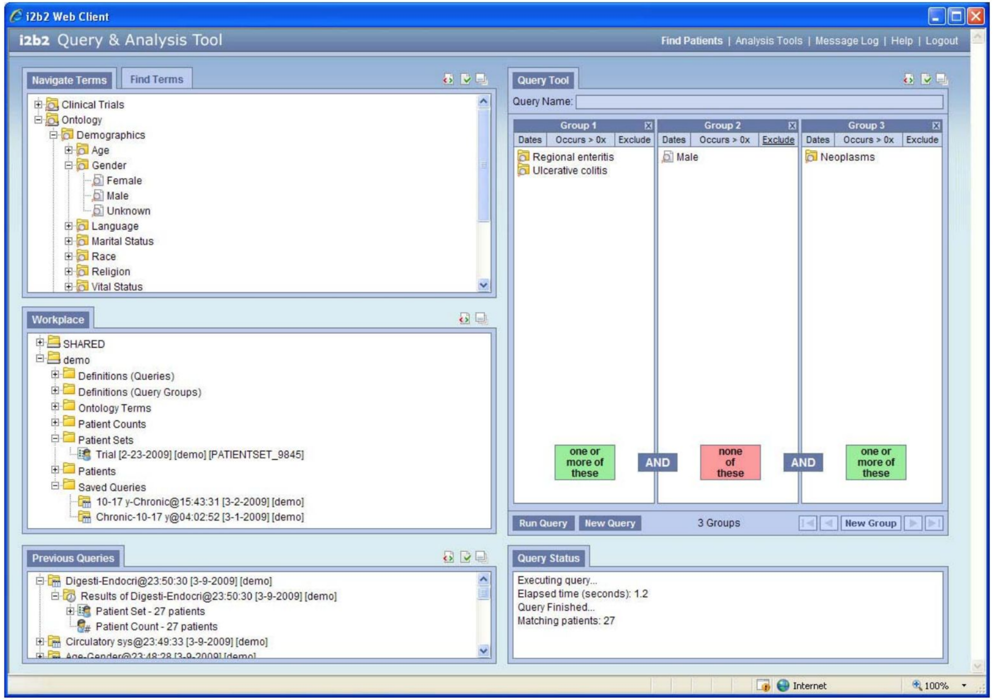
Real World Data

- Very heterogeneous, especially across hospitals / institutes
- Standardization
  - Ontology / coding
  - Format
- Platform
  - Routine hospital use
  - Data warehouse
  - ***i2b2 star schema [Murphy 2010]***
  - Research cohort query interface



Real World Data

- Very heterogeneous, especially across hospitals / institutes
- Standardization
  - ***Ontology / coding***
  - Format
- Platform
  - Routine hospital use
  - Data warehouse
  - i2b2 star schema [Murphy 2010]
  - ***Research cohort query interface***



Why Now?

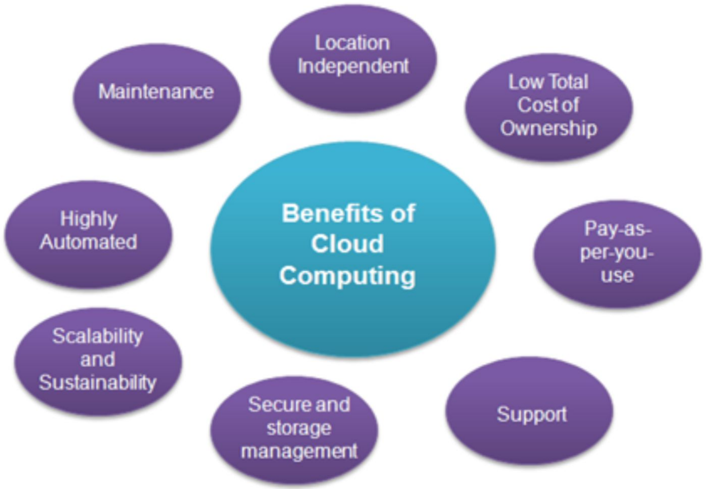
- High quality (?) multimodal data
  - Variety, volume, velocity
  - Social media, sensor, wearable, vital signs, lab data, notes, imaging, -omics
- Adoption of EHR
- Standardization
  - UMLS, SNOMED, ICD, RxNorm, LOINC, ...
  - HL7, FHIR, ...
- Advances and democratization in ML
- ***Open-source / algorithms & tools***
- Computing resources

Why Now?

- High quality (?) multimodal data
  - Variety, volume, velocity
  - Social media, sensor, wearable, vital signs, lab data, notes, imaging, -omics
- Adoption of EHR
- Standardization
  - UMLS, SNOMED, ICD, RxNorm, LOINC, ...
  - HL7, FHIR, ...
- Advances and democratization in ML
- Open-source / algorithms & tools
- ***Computing resources***

Computing Resources

- For education / research
  - Google Cloud
    - TPU
  - Amazon AWS
    - EC2 for computing
    - S3 for storage
- IRB security level / HIPAA for privacy
  - Level 5 - Extremely sensitive information
  - ***Level 4 - Very sensitive information***
  - Level 3 - Sensitive, or Confidential information
  - Level 2 - Benign information to be held confidentially
  - Level 1 - Non-confidential research information
- AWS educate



The 18 Identifiers defined by HIPAA are:

- ☑ Name
- ☑ Postal address
- ☑ All elements of dates except year
- ☑ Telephone number
- ☑ Fax number
- ☑ Email address
- ☑ URL address
- ☑ IP address
- ☑ Social security number
- ☑ Account numbers
- ☑ License numbers
- ☑ Medical record number
- ☑ Health plan beneficiary #
- ☑ Device identifiers and their serial numbers
- ☑ Vehicle identifiers and serial number
- ☑ Biometric identifiers (finger and voice prints)
- ☑ Full face photos and other comparable images
- ☑ Any other unique identifying number, code, or characteristic

Why Now?

- High quality (?) multimodal data
  - Variety, volume, velocity
  - Social media, sensor, wearable, vital signs, lab data, notes, imaging, -omics
- Adoption of EHR
- Standardization
  - UMLS, SNOMED, ICD, RxNorm, LOINC, ...
  - HL7, FHIR, ...
- ***Advances and democratization in ML (modeling!)***
- Open-source / algorithms & tools
- Computing resources

Sciences Behind

Data organization / structuring

- Database
- Knowledge representation / Ontology
- Visualization

Learning from data (algorithms, machine learning)

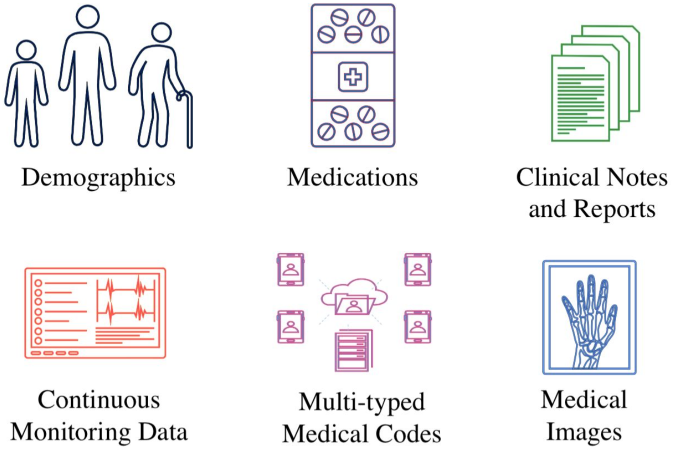
- Statistics, vector calculus, probability, optimization, information theory, …
- Computer vision, natural language processing, signal processing, …

Why ML?

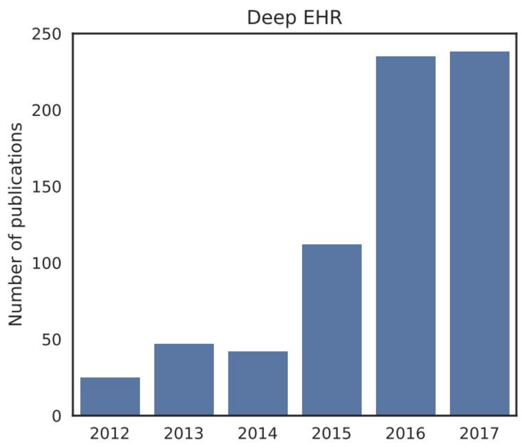
- Failure (?) of classical symbolic AI
    - A lot of knowledge is intuitive, difficult to put in rules and facts, not consciously accessible
- Principles giving rise to intelligence via learning
- To facilitate learning higher levels of abstraction [Bengio 2013, LeCun 2015]
    - Deep learning
- Can be described compactly since our intelligence is not just the result of a huge bag of tricks and pieces of knowledge, but of general mechanisms to acquire knowledge [Bengio 2013]
- ***Get knowledge directly from data and experience → learn from data***

Why ML/DL for EHR?

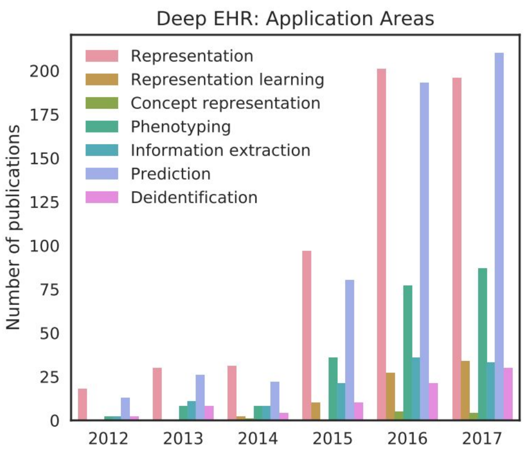
- Demographics, diagnoses, laboratory test results, medication prescriptions, clinical notes, medical images, ...
- Challenging
  - data quality (noisy, biased, ...)
  - data and annotation availability
  - heterogeneity of data types
- ***Traditional modeling → feature engineering***
  - labor intensive efforts
  - expert-defined phenotyping
  - ad-hoc feature engineering
  - limited generalizability across datasets or institutions
- ***Deep learning → learning hidden representations***
  - expert-driven feature engineering to data-driven feature construction



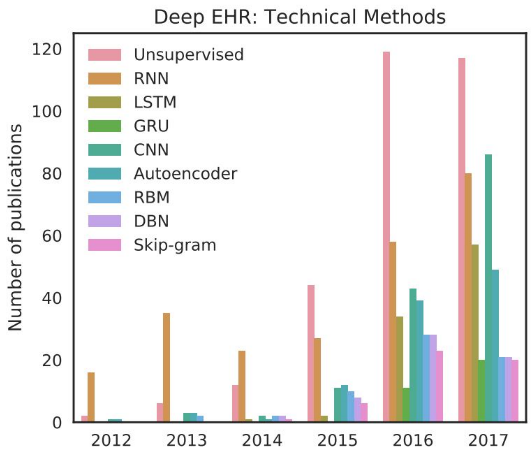
Deep! (2017/06)



Deep EHR
Number of publications
0
50
100
150
200
250
2012
2013
2014
2015
2016
2017



Deep EHR: Application Areas
Number of publications
Representation
Representation learning
Concept representation
Phenotyping
Information extraction
Prediction
Deidentification
0
25
50
75
100
125
150
175
200
2012
2013
2014
2015
2016
2017



Deep EHR: Technical Methods
Number of publications
Unsupervised
RNN
LSTM
GRU
CNN
Autoencoder
RBM
DBN
Skip-gram
0
20
40
60
80
100
120
2012
2013
2014
2015
2016
2017

How?

- How to develop machine learning models for healthcare [Chen 2019]
- Machine Learning for Clinical Predictive Analytics [Weng 2019]

- ***Problem definition***
- ***Data curation***
- ***ML model development***
- ***Validation***
- ***Assessment of clinical impact***
- ***(Deployment and monitoring)***

More details → Oct 29 Workshop!

comment

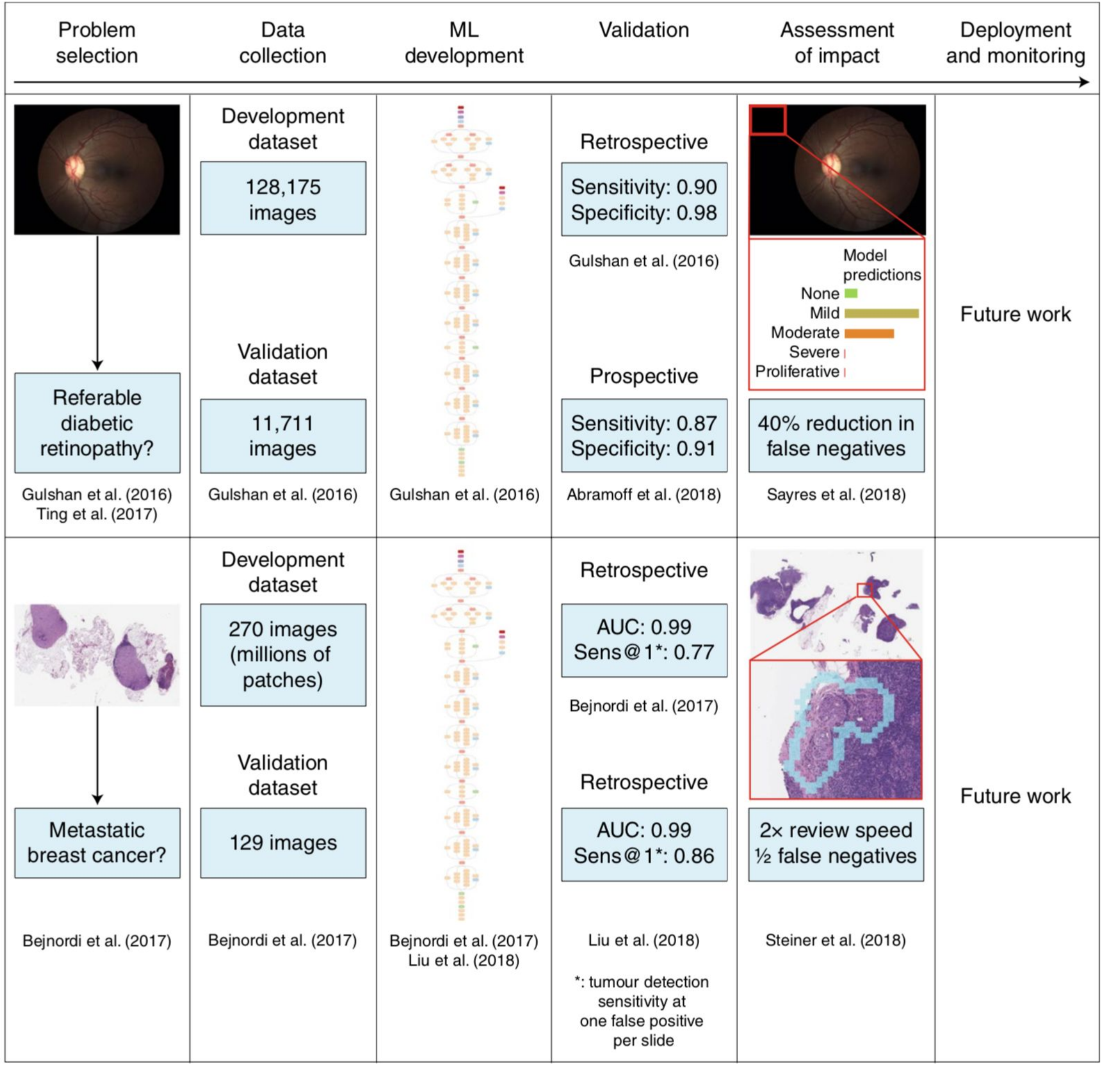
How to develop machine learning models for healthcare

Rapid progress in machine learning is enabling opportunities for improved clinical decision support. Importantly, however, developing, validating and implementing machine learning models for healthcare entail some particular considerations to increase the chances of eventually improving patient care.

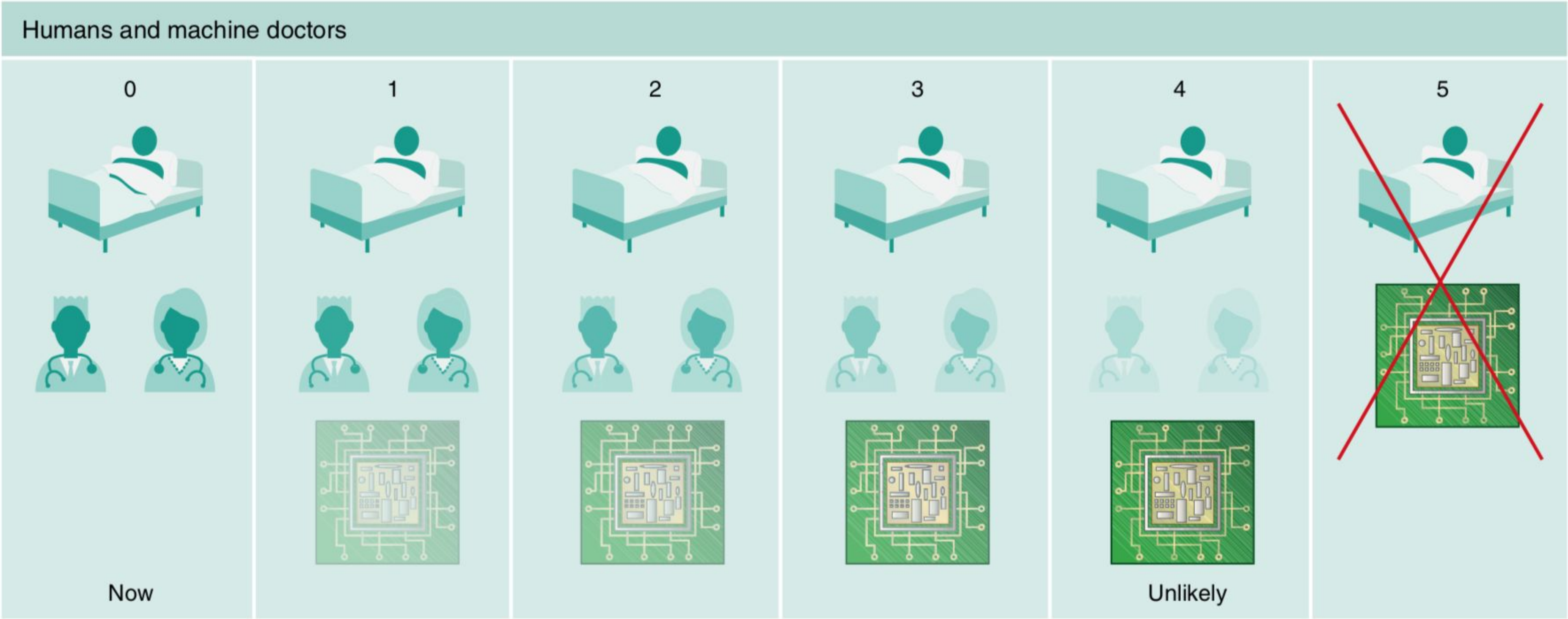
Po-Hsuan Cameron Chen, Yun Liu and Lily Peng

Machine Learning for Clinical Predictive Analytics

Wei-Hung Weng [1]


Problem selection
Data collection
ML development
Validation
Assessment of impact
Deployment and monitoring
Development dataset
128,175 images
Retrospective
Sensitivity: 0.90
Specificity: 0.98
Gulshan et al. (2016)
Model predictions
None
Mild
Moderate
Severe
Proliferative
Future work
Validation dataset
Prospective
Referable diabetic retinopathy?
11,711 images
Sensitivity: 0.87
Specificity: 0.91
40% reduction in false negatives
Gulshan et al. (2016)
Ting et al. (2017)
Gulshan et al. (2016)
Gulshan et al. (2016)
Abramoff et al. (2018)
Sayres et al. (2018)
Development dataset
Retrospective
270 images (millions of patches)
AUC: 0.99
Sens@1*: 0.77
Bejnordi et al. (2017)
Validation dataset
Retrospective
Future work
Metastatic breast cancer?
129 images
AUC: 0.99
Sens@1*: 0.86
2× review speed
½ false negatives
Bejnordi et al. (2017)
Bejnordi et al. (2017)
Bejnordi et al. (2017)
Liu et al. (2018)
Liu et al. (2018)
Steiner et al. (2018)
*: tumour detection sensitivity at one false positive per slide

GIGO!



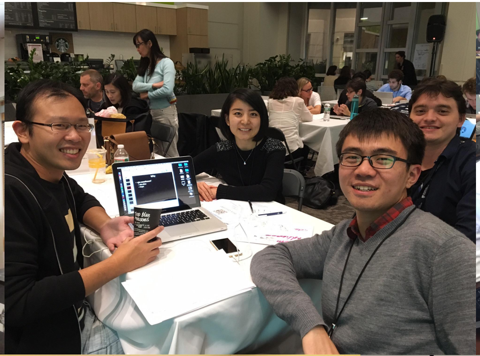
Topol. Nat Med. 2019

Interpretability

https://xkcd.com/1838/

Collaborative Medicine

- Courses, Workshops, Symposium, Datathon
- Japan, China, Singapore, Philippines, Denmark, ***Bhutan***



Collaborative Medicine

- Collaboration > Expert or AI only [Steiner 2018]
    - Expert: high-level integration, interpretation and decision making (learn Bayesian logic, statistics, and data science and be aware of other sources of information!)
    - AI: pattern recognition and massive repetitive tasks
- Sharing
    - AI researchers open-source platforms and algorithms
    - You can open-source data and knowledge
    - We can do nothing without sharing!
- Experience
    - Learning from AI
    - Communicating with AI

Further Readings

- Theory and mathematics
  - Coursera Machine Learning (Andrew Ng)
  - Deep learning book
  - PMRL
  - Stanford CS224n: Natural Language Processing with Deep Learning
  - Stanford CS231n: Convolutional Neural Networks for Visual Recognition
- Practical
  - TensorFlow, PyTorch, Keras, Scikit-learn documents, guide, tutorials
  - Google Machine Learning Crash Course
  - Do more projects!

Seeking Papers

- Clinical-driven (journals)
  - JAMA
  - Nature Medicine
  - npj Digital Medicine
  - JAMIA
  - AMIA (conference)
- ML/AI-driven (conferences / workshops)
  - MLHC
  - NeurIPS ML4H Workshop
  - NAACL Clinical NLP Workshop
  - MICCAI
  - NeurIPS, ICML, ICLR, KDD, AAAI, ACL, NAACL, CVPR, ...

tl;dr

- ***Good problems + reliable data (good infra!) >>> fancy ML algorithms***
- Collaboration is the key
  - Collaboration > clinical experts or engineers/AI only [Steiner 2018]
  - Sharing with your colleagues and collaborators
- Precision Medicine
  - Large multimodal data available, standardization, computing resources and open-source algorithms / tools
- Model development [Chen 2019, Weng 2019]
  - Problem definition → data curation → model development → validation → assessment of clinical impact → deployment and monitoring
- Wei-Hung Weng [`ckbjimmy@mit.edu`]