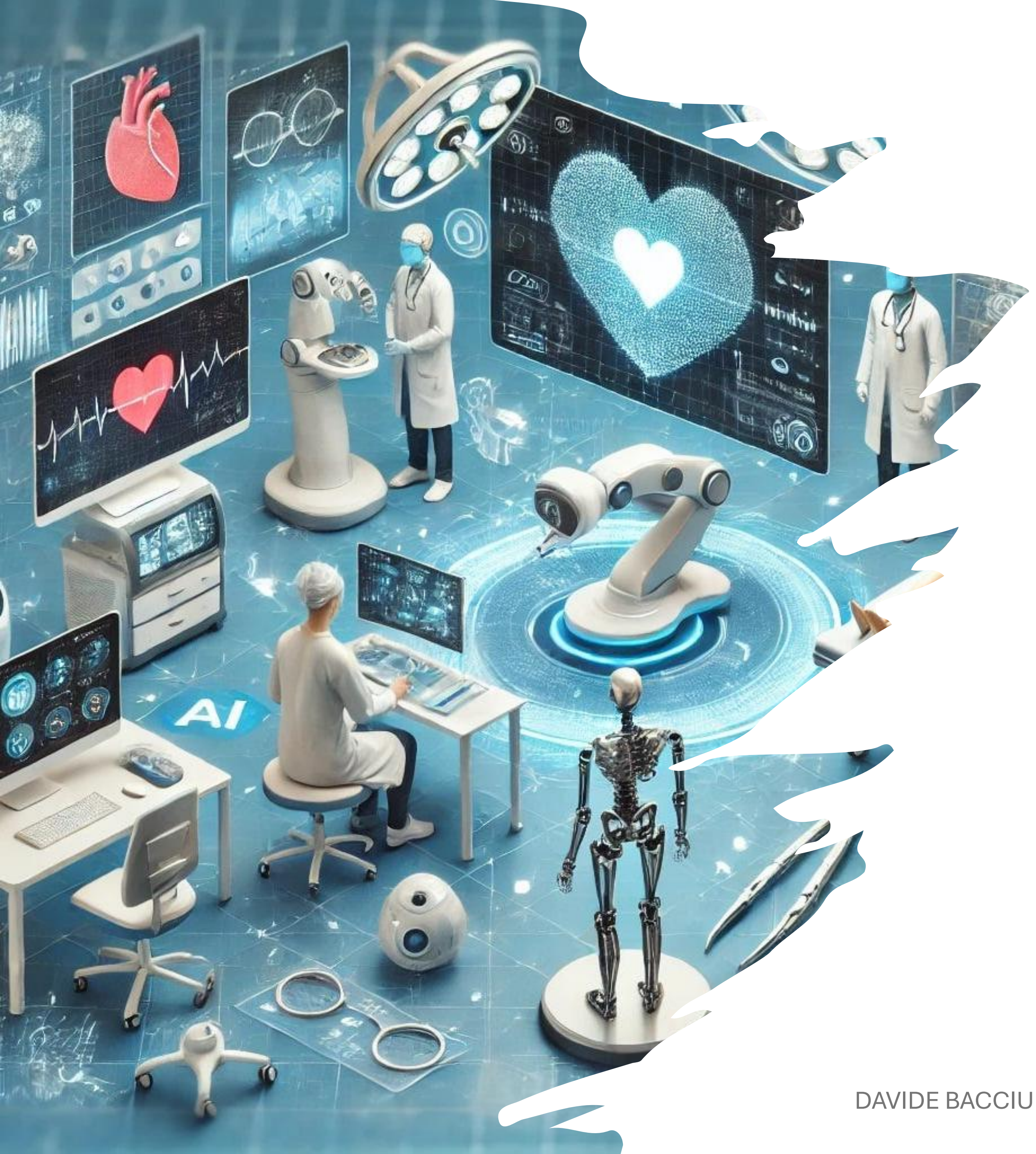
Artificial Intelligence for Digital Health – Course Overview

Artificial Intelligence for Digital Health (AID)
M.Sc. in Digital Health – University of Pisa
Davide Bacciu (davide.bacciu@unipi.it)
Marco Podda (marco.podda@unipi.it)

Lecture Outline

- Course overview
- A history of Artificial Intelligence in healthcare and life sciences
- Perks of biomedical/health data
- Course execution and examination

Course Overview



Course objectives

Train ICT specialists who

- Understand the challenges of working with health data, including physiological time-series, clinical text, and medical imaging
- Know the leading AI-based methodologies and understand how to apply them to relevant digital health problems
- Have practical insights into the main AI technological solutions and programming resources and how to use them in digital health
- Understand how to robustly use AI methodologies and develop trusted AI-based digital health applications

Prerequisites

Foundational knowledge of

- Calculus and linear algebra (Appendix B, Prince textbook)
- Elements of probability and statistics (Appendix C, Prince textbook)
- Basics of optimization (compendium)

Programming experience with Python is assumed for practical labs

- Previous experience with AI/ML Python libraries is a plus

Organization of Course Contents

- **Preliminaries**
  - Course overview
  - Probability and statistics refresher for AI
- **Fundamentals of AI in healthcare**
  - Machine learning basics and baselines (regression, neural networks)
  - Risk stratification and survival analysis
  - Probabilistic/Bayesian models and causality
- **Deep Learning**
  - Deep learning basics and baselines
  - AI for medical imaging (classification, segmentation, recognition)
  - AI for sequential data in healthcare
  - AI for clinical text and language models for health
- **Advanced topics**
  - Graphs in Health and Life Sciences
  - AI applications in healthcare
  - Research directions and challenges in AI for health
  - Guest lectures

Hands-on laboratories will complement methodological lectures

Course Instructors

Davide Bacciu (Lecturer) - davide.bacciu@unipi.it

- Professor of Machine Learning - Head of Pervasive AI lab
- Lecturer of deep and generative learning at the M.Sc. in AI
- Startup: quantitative brain techs, AI for harsh environments, AI for regulatory compliance
- Office hours: on request, via mail

Marco Podda (Laboratory Lecturer)

- Researcher, Computer Science Department
- Learning from graph data and generative models
- Applications in healthcare domain
- Office hours: on request, via mail

Dates, times and format

Day	Time
Tuesday – Room C	16.00-18.00
Wednesday – Room L1	11.00-13.00
Thursday – Room L1	16.00-18.00

- Course comprises ~24 lectures and ~12 labs
- Course lectures and labs will be given in-person (no streaming)
- Video recording of the lectures will be made available on the shared Teams drive (best effort, no guarantees)

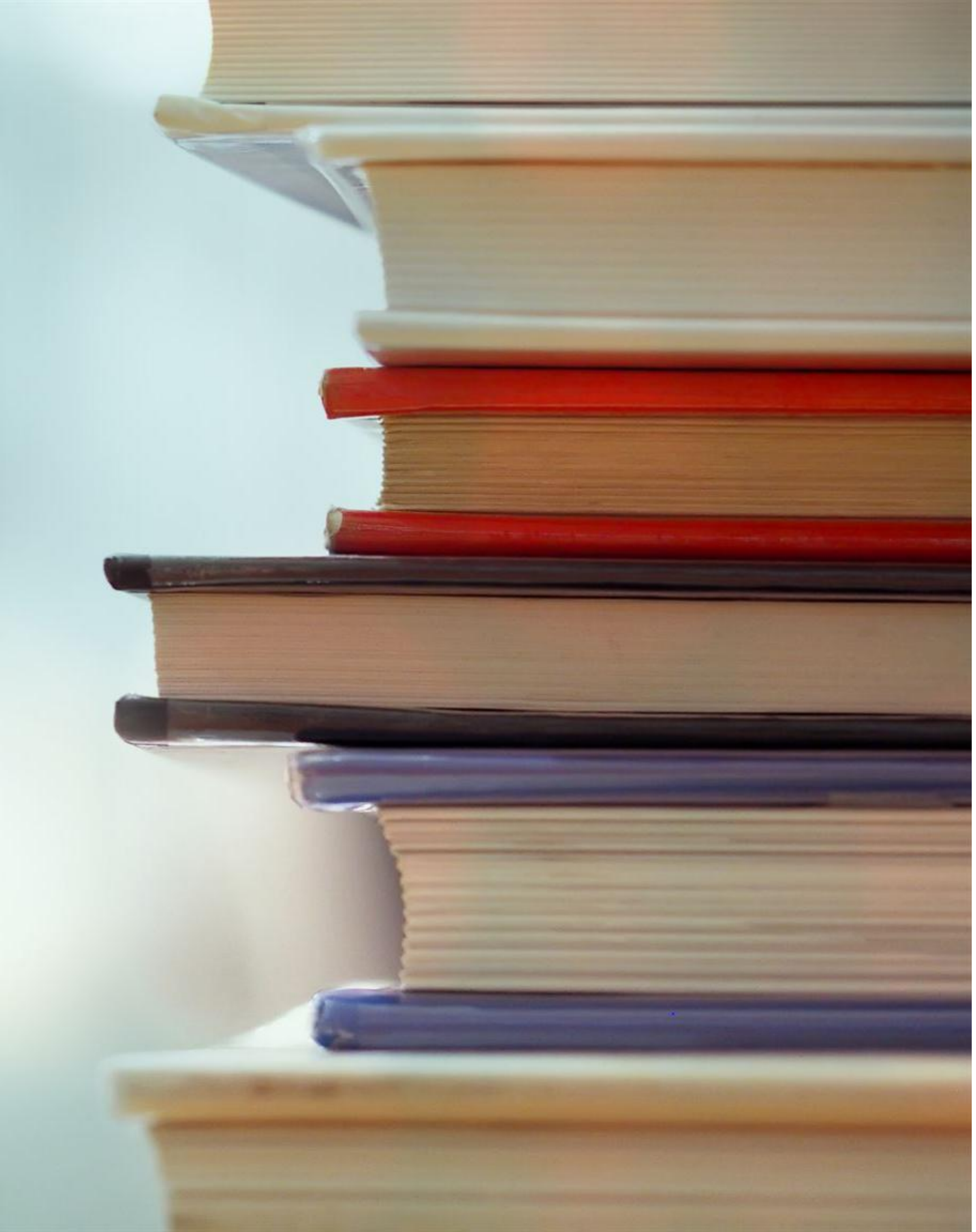
Resources

Reference Webpage on Moodle:
https://elearning.di.unipi.it/course/view.php?id=1024

Here you will find

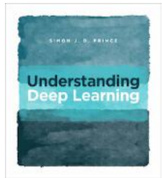
- Course information and up-to-date news
- Lecture slides
- Articles and course materials
- Project assignments and exams materials

Make sure to **subscribe** to the course on Moodle



Reference Books

- Reference for foundational AI/ML knowledge



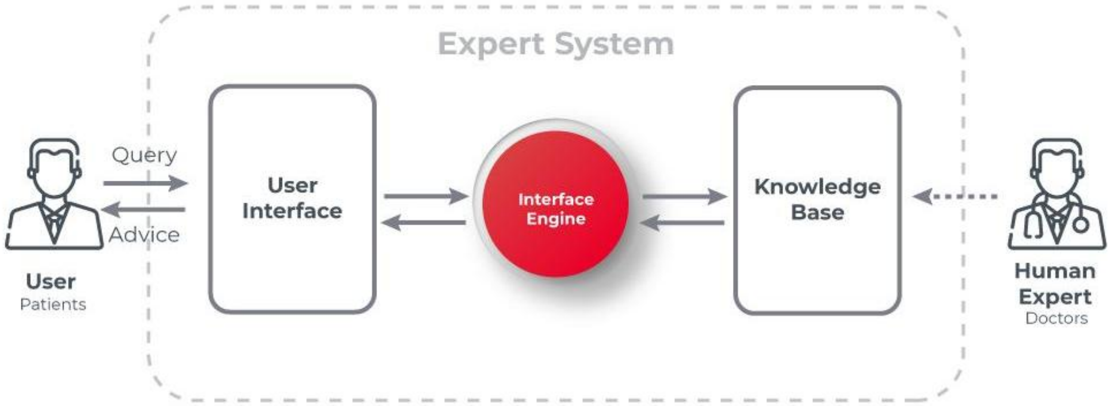
  - Simon J.D. Prince, Understanding Deep Learning, MIT Press (2023)
  - Additional materials on the book website
- Reference for AI in healthcare
  - G.J. Simon, C. Aliferis, AI and ML in health Care and Medical Sciences, Springer, 2024
- Both available open-source

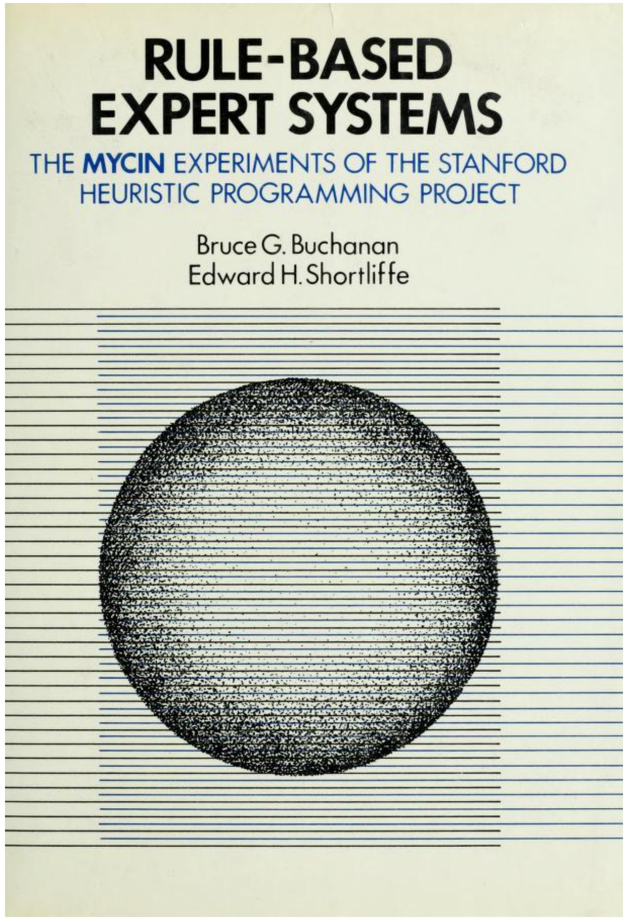
AI in health and life sciences

AI in healthcare: a long-standing quest

The expert systems era (1970-1980)

- MYCIN: identify bacteria causing severe infections a proposing a therapy
- Rule-based (600) + simple inference engine



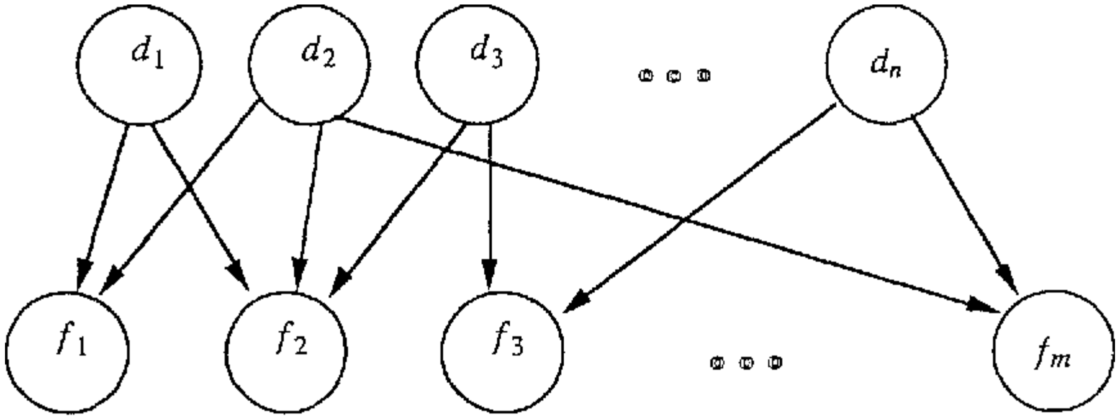


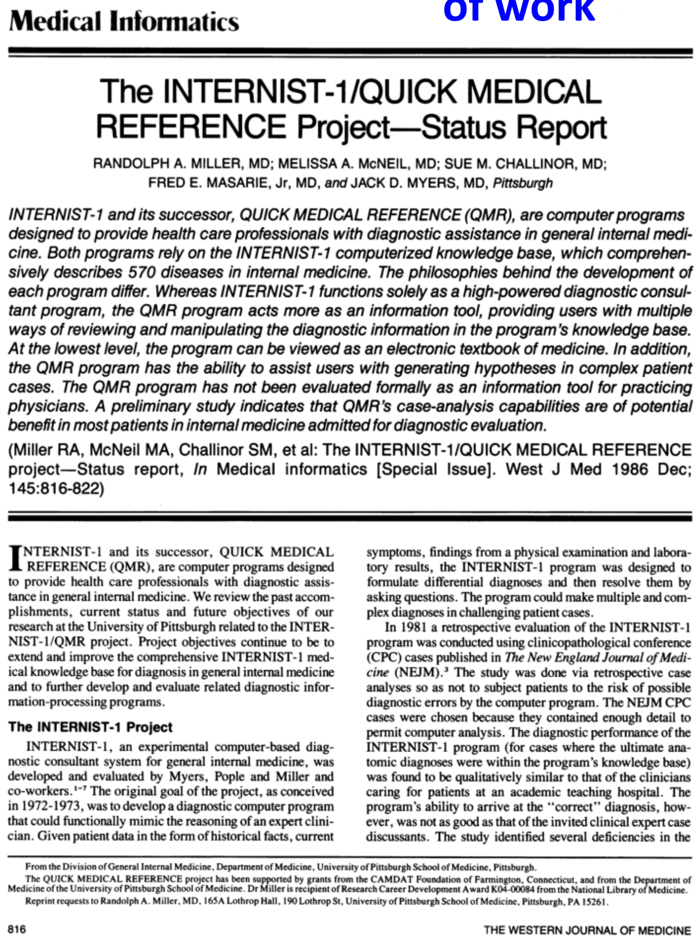
1980s: The rise of Bayesian methods

15 person-years of work

INTERNIST-1/Quick Medical Reference

- Introduced decision trees and belief networks
- The fundamentals of Bayesian networks and approximated inference





Medical Informatics

The INTERNIST-1/QUICK MEDICAL REFERENCE Project—Status Report

RANDOLPH A. MILLER, MD; MELISSA A. McNEIL, MD; SUE M. CHALLINOR, MD; FRED E. MASARIE, Jr, MD, *and* JACK D. MYERS, MD, *Pittsburgh*

INTERNIST-1 and its successor, QUICK MEDICAL REFERENCE (QMR), are computer programs designed to provide health care professionals with diagnostic assistance in general internal medicine. Both programs rely on the INTERNIST-1 computerized knowledge base, which comprehensively describes 570 diseases in internal medicine. The philosophies behind the development of each program differ. Whereas INTERNIST-1 functions solely as a high-powered diagnostic consultant program, the QMR program acts more as an information tool, providing users with multiple ways of reviewing and manipulating the diagnostic information in the program's knowledge base. At the lowest level, the program can be viewed as an electronic textbook of medicine. In addition, the QMR program has the ability to assist users with generating hypotheses in complex patient cases. The QMR program has not been evaluated formally as an information tool for practicing physicians. A preliminary study indicates that QMR's case-analysis capabilities are of potential benefit in most patients in internal medicine admitted for diagnostic evaluation.

(Miller RA, McNeil MA, Challinor SM, et al: The INTERNIST-1/QUICK MEDICAL REFERENCE project—Status report, *In* Medical informatics [Special Issue]. West J Med 1986 Dec; 145:816-822)

INTERNIST-1 and its successor, QUICK MEDICAL REFERENCE (QMR), are computer programs designed to provide health care professionals with diagnostic assistance in general internal medicine. We review the past accomplishments, current status and future objectives of our research at the University of Pittsburgh related to the INTERNIST-1/QMR project. Project objectives continue to be to extend and improve the comprehensive INTERNIST-1 medical knowledge base for diagnosis in general internal medicine and to further develop and evaluate related diagnostic information-processing programs.

The INTERNIST-1 Project

INTERNIST-1, an experimental computer-based diagnostic consultant system for general internal medicine, was developed and evaluated by Myers, Pople and Miller and co-workers.[1-7] The original goal of the project, as conceived in 1972-1973, was to develop a diagnostic computer program that could functionally mimic the reasoning of an expert clinician. Given patient data in the form of historical facts, current symptoms, findings from a physical examination and laboratory results, the INTERNIST-1 program was designed to formulate differential diagnoses and then resolve them by asking questions. The program could make multiple and complex diagnoses in challenging patient cases.

In 1981 a retrospective evaluation of the INTERNIST-1 program was conducted using clinicopathological conference (CPC) cases published in *The New England Journal of Medicine* (NEJM).[3] The study was done via retrospective case analyses so as not to subject patients to the risk of possible diagnostic errors by the computer program. The NEJM CPC cases were chosen because they contained enough detail to permit computer analysis. The diagnostic performance of the INTERNIST-1 program (for cases where the ultimate anatomic diagnoses were within the program's knowledge base) was found to be qualitatively similar to that of the clinicians caring for patients at an academic teaching hospital. The program's ability to arrive at the "correct" diagnosis, however, was not as good as that of the invited clinical expert case discussants. The study identified several deficiencies in the

From the Division of General Internal Medicine, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh.

The QUICK MEDICAL REFERENCE project has been supported by grants from the CAMDAT Foundation of Farmington, Connecticut, and from the Department of Medicine of the University of Pittsburgh School of Medicine. Dr Miller is recipient of Research Career Development Award K04-00084 from the National Library of Medicine.

Reprint requests to Randolph A. Miller, MD, 165A Lothrop Hall, 190 Lothrop St, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261.

816 THE WESTERN JOURNAL OF MEDICINE

1990s: Neural networks enter medicine

- Nearly 100 studies on neural networks on clinical data in 1990
- Networks with few selected and hand engineered inputs
- Issues:
    - Did not fit well into clinical workflow
    - Hard to source enough training data
    - Poor generalization

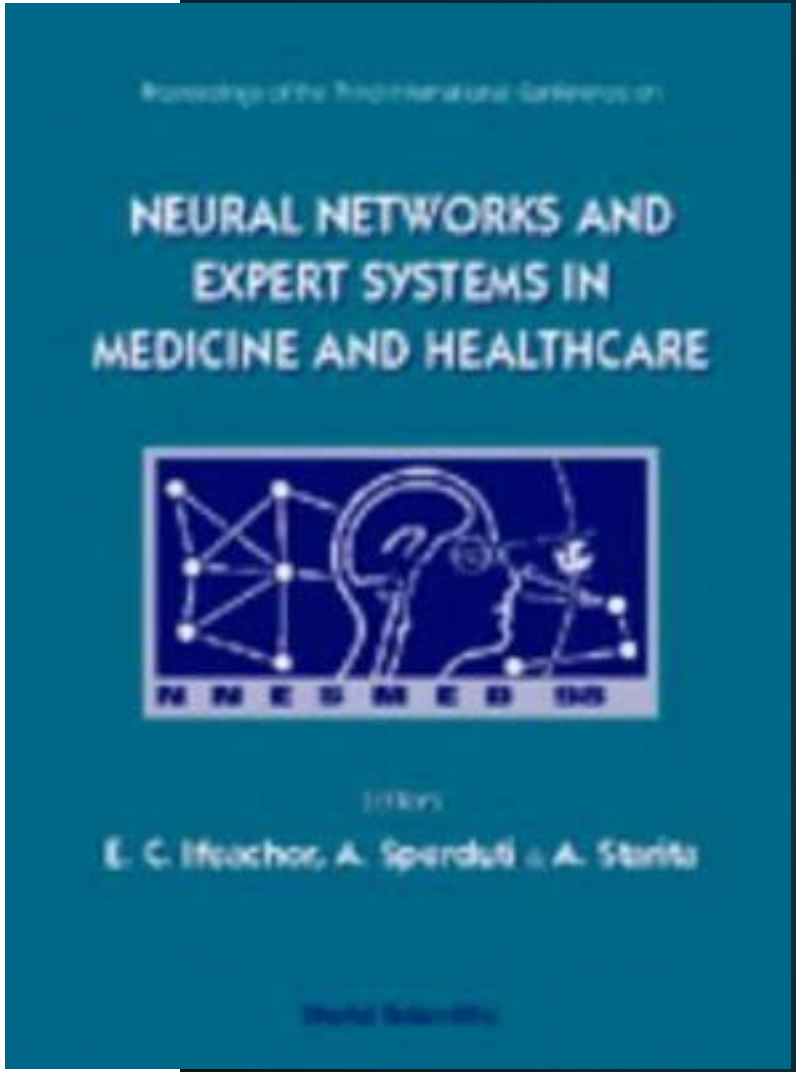
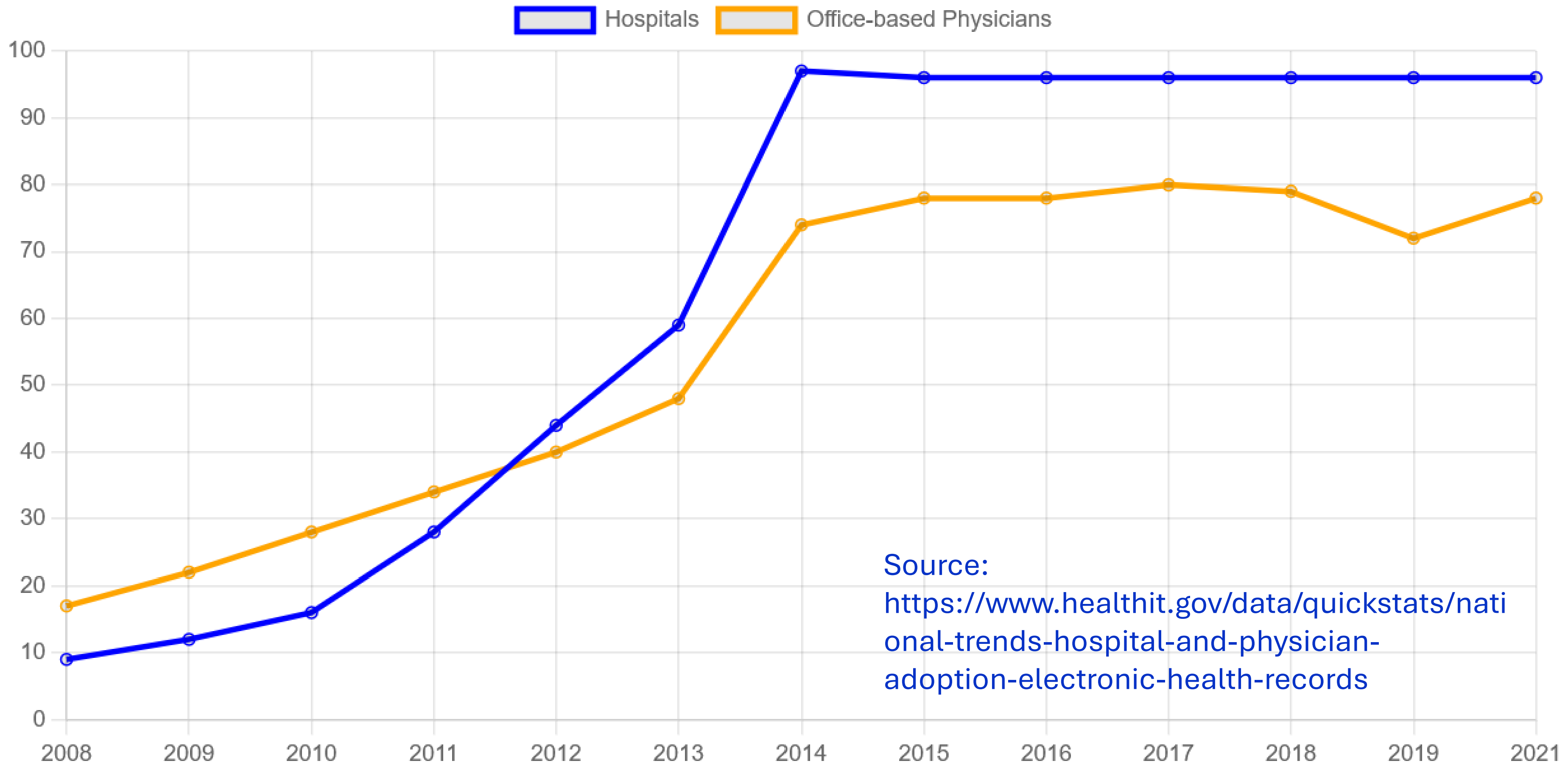


Table 1 • 25 Neural Network Studies in Medical Decision Making*

Subject	No. of Examples		P†	Network	D‡	Accuracy§	
	Training	Test				Neural	Other
Breast cancer[4]	57	20	60	9–15–2	0.6	80	75
Vasculitis[2]	404	403	73	8–5–1	8.0	94	—
Myocardial infarction[6]	351	331	89	20–10–10–1	1.1	97	84
Myocardial infarction[8]	356	350	87	20–10–10–1	1.1	97	94
Low back pain[11]	100	100	25	50–48–2	0.2	90	90
Cancer outcome[13]	5,169	3,102	—	54–40–1	1.4	0.779	0.776
Psychiatric length of stay[17]	957	106	73	48–400–4	0.2	74	76
Intensive care outcome[23]	284	138	91	27–18–1	0.5	0.82	0.82
Skin tumor[21]	150	100	80	18	—	80	90
Evoked potentials[35]	100	67	52	14–4–3	3.8	77	77
Head injury[47]	500	500	50	6–3–3	20	66	77
Psychiatric outcome[54]	289	92	60	41–10–1	0.7	79	—
Tumor classification[55]	53	6	38	8–9–3	1.4	99	88
Dementia[57]	75	18	19	80–10–7–7	0.6	61	—
Pulmonary embolism[59]	607	606	69	50–4–1	2.9	0.82	0.83
Heart disease[62]	460	230	54	35–16–8–2	3	83	84
Thyroid function[62]	3,600	1,800	93	21–16–8–3	22	98	93
Breast cancer[62]	350	175	66	9–4–4–2	10	97	96
Diabetes[62]	384	192	65	8–4–4–2	12	77	75
Mycardial infarction[63]	2,856	1,429	56	291–1	9.8	85	—
Hepatitis[65]	39	42	38	4–4–3	3.3	74	79
Psychiatric admission[76]	319	339	85	53–1–1	6.0	91	—
Cardiac length of stay[83]	713	696	73	15–12–1	3.5	0.70	—
Anti-cancer agents[89]	127	141	25	60–7–6	1.5	91	86
Ovarian cancer[91]	75	98	—	6–6–2	2.6	84	81
MEDIAN VALUE	350	175	71	20	2.8		

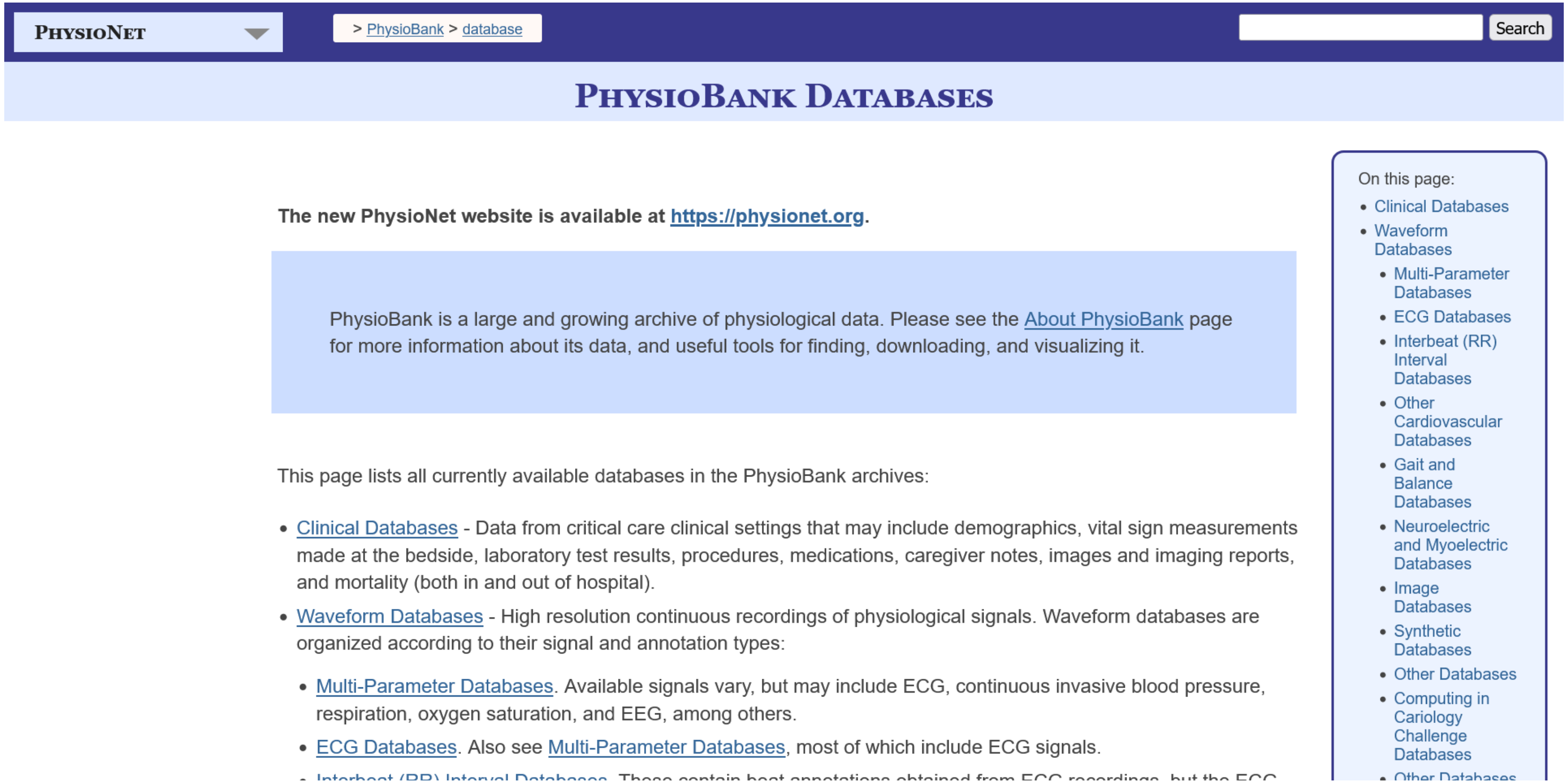
Then things started taking a favourable path

Adoption of Electronic Health Records (EHRs) began a steady growth



Source: https://www.healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records

Availability of biobanks and resources



PhysioNet

> PhysioBank > database

Search

PhysioBank Databases

The new PhysioNet website is available at https://physionet.org.

PhysioBank is a large and growing archive of physiological data. Please see the About PhysioBank page for more information about its data, and useful tools for finding, downloading, and visualizing it.

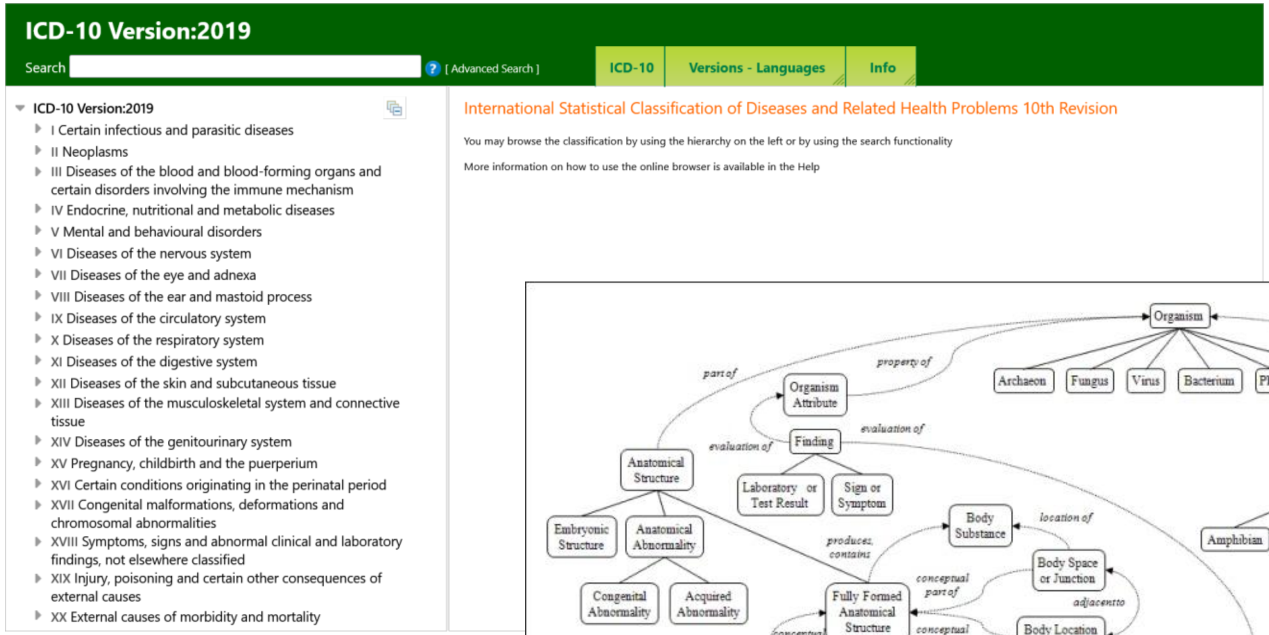
This page lists all currently available databases in the PhysioBank archives:

- Clinical Databases - Data from critical care clinical settings that may include demographics, vital sign measurements made at the bedside, laboratory test results, procedures, medications, caregiver notes, images and imaging reports, and mortality (both in and out of hospital).
- Waveform Databases - High resolution continuous recordings of physiological signals. Waveform databases are organized according to their signal and annotation types:
  - Multi-Parameter Databases. Available signals vary, but may include ECG, continuous invasive blood pressure, respiration, oxygen saturation, and EEG, among others.
  - ECG Databases. Also see Multi-Parameter Databases, most of which include ECG signals.

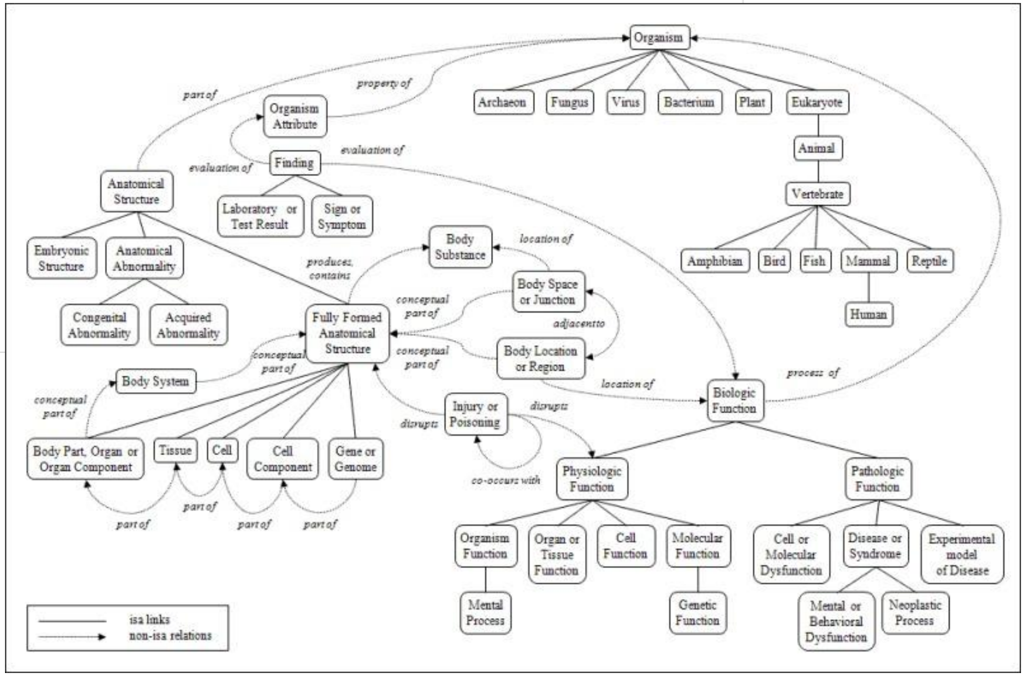
On this page:

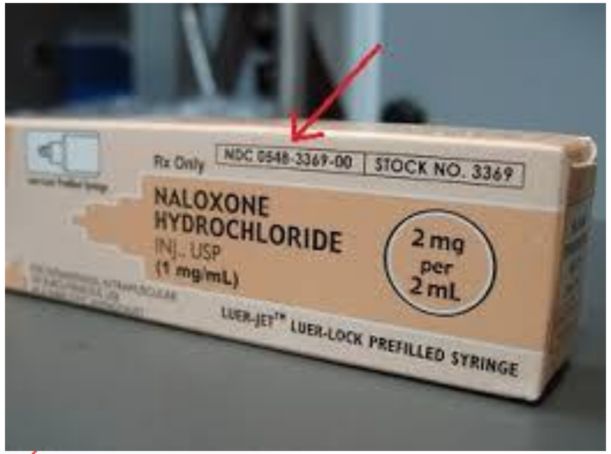
- Clinical Databases
- Waveform Databases
  - Multi-Parameter Databases
  - ECG Databases
  - Interbeat (RR) Interval Databases
  - Other Cardiovascular Databases
  - Gait and Balance Databases
  - Neuroelectric and Myoelectric Databases
  - Image Databases
  - Synthetic Databases
  - Other Databases
  - Computing in Cariology Challenge Databases

Standardization



ICD: Classification of Diseases



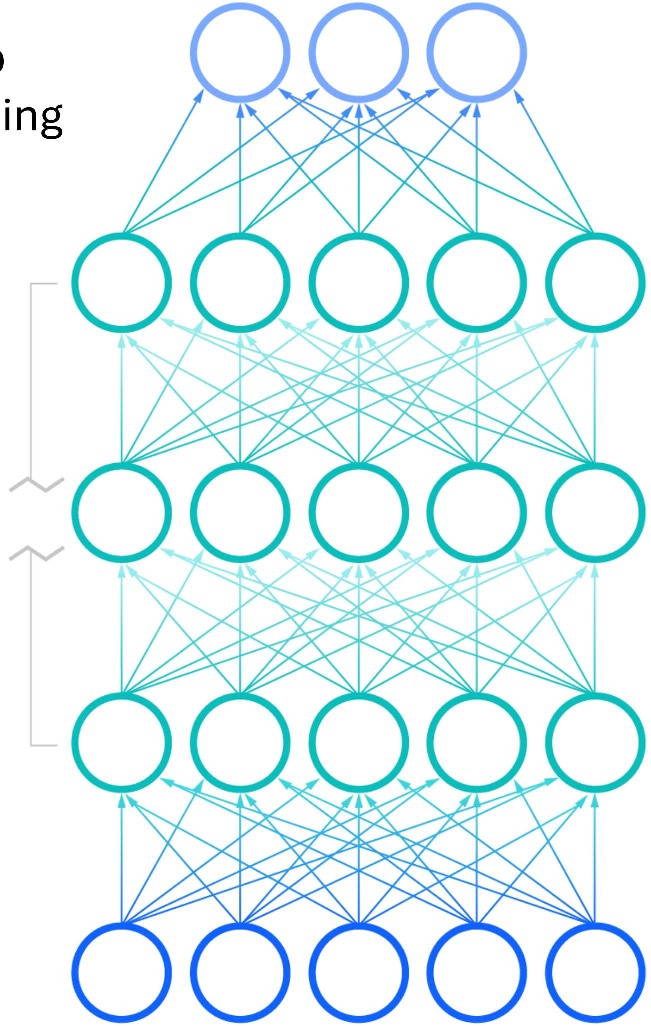


National drug codes

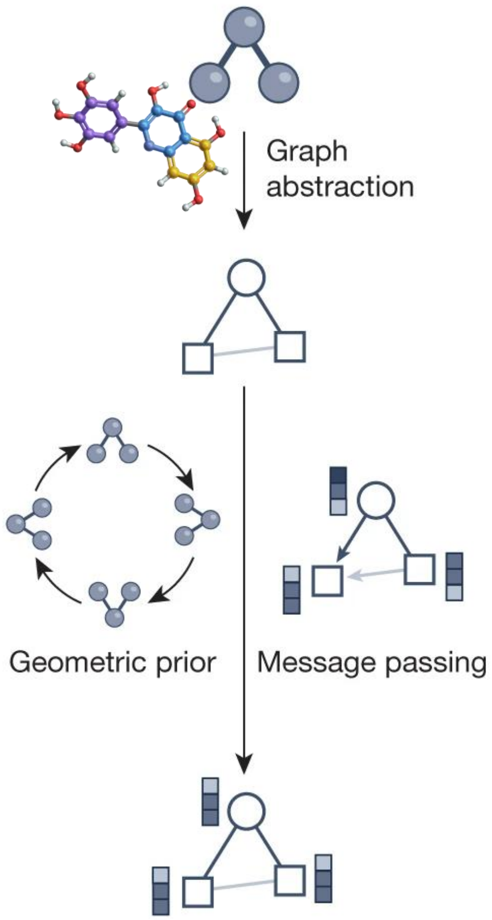
Unified Medical Language System: millions of medical concepts

Key AI advances

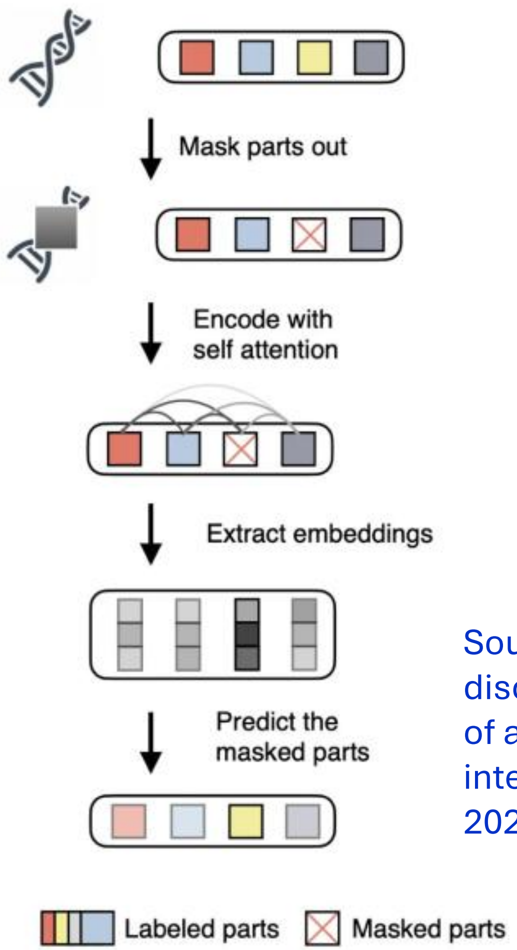
Deep learning



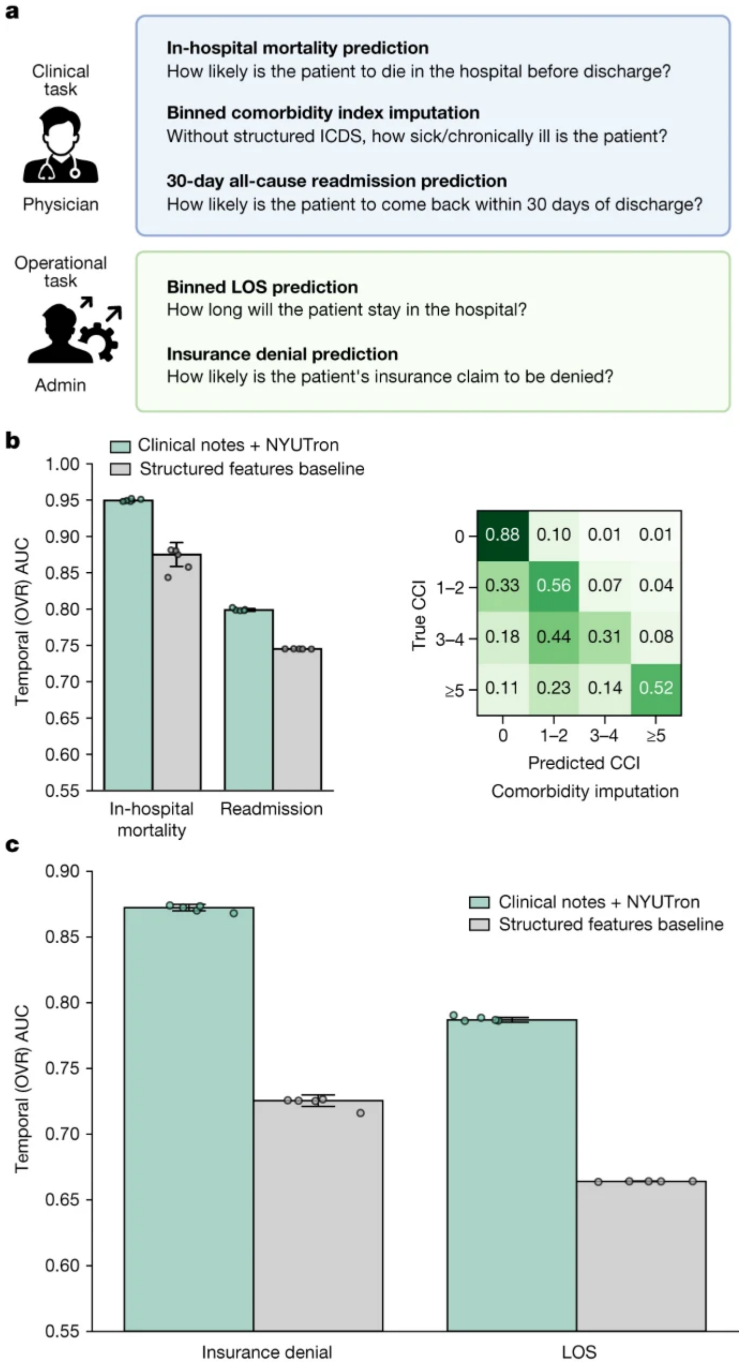
Learning from complex data



Self-supervised learning

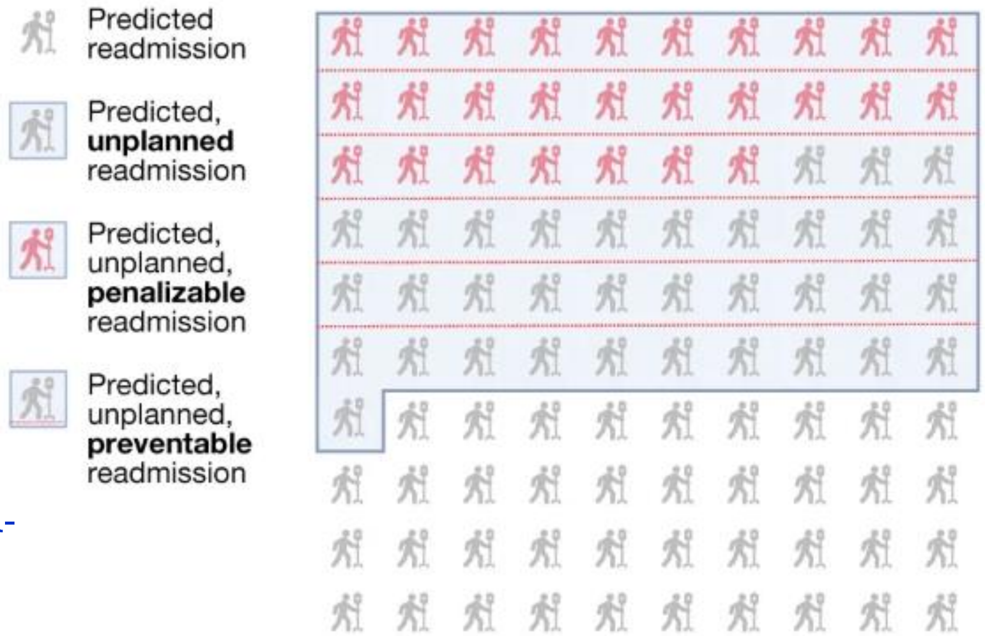


Source: Scientific discovery in the age of artificial intelligence, Nature 2023



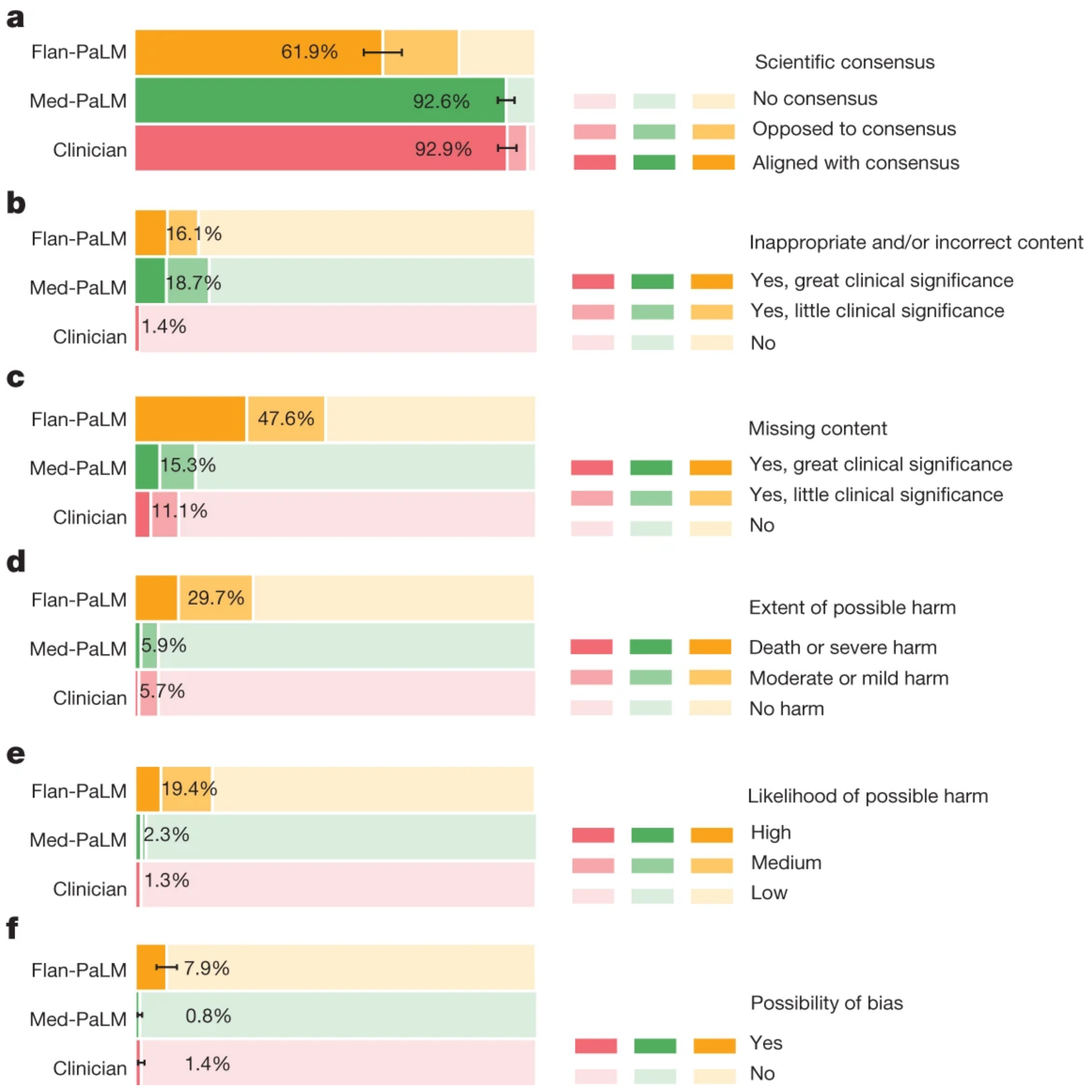
AI in Digital Health in the LLM era

Assessment of an LLM-based agent (NYUTron) for multiple clinical and operational healthcare tasks



Health system-scale language models are all-purpose prediction engines, Nature 2023

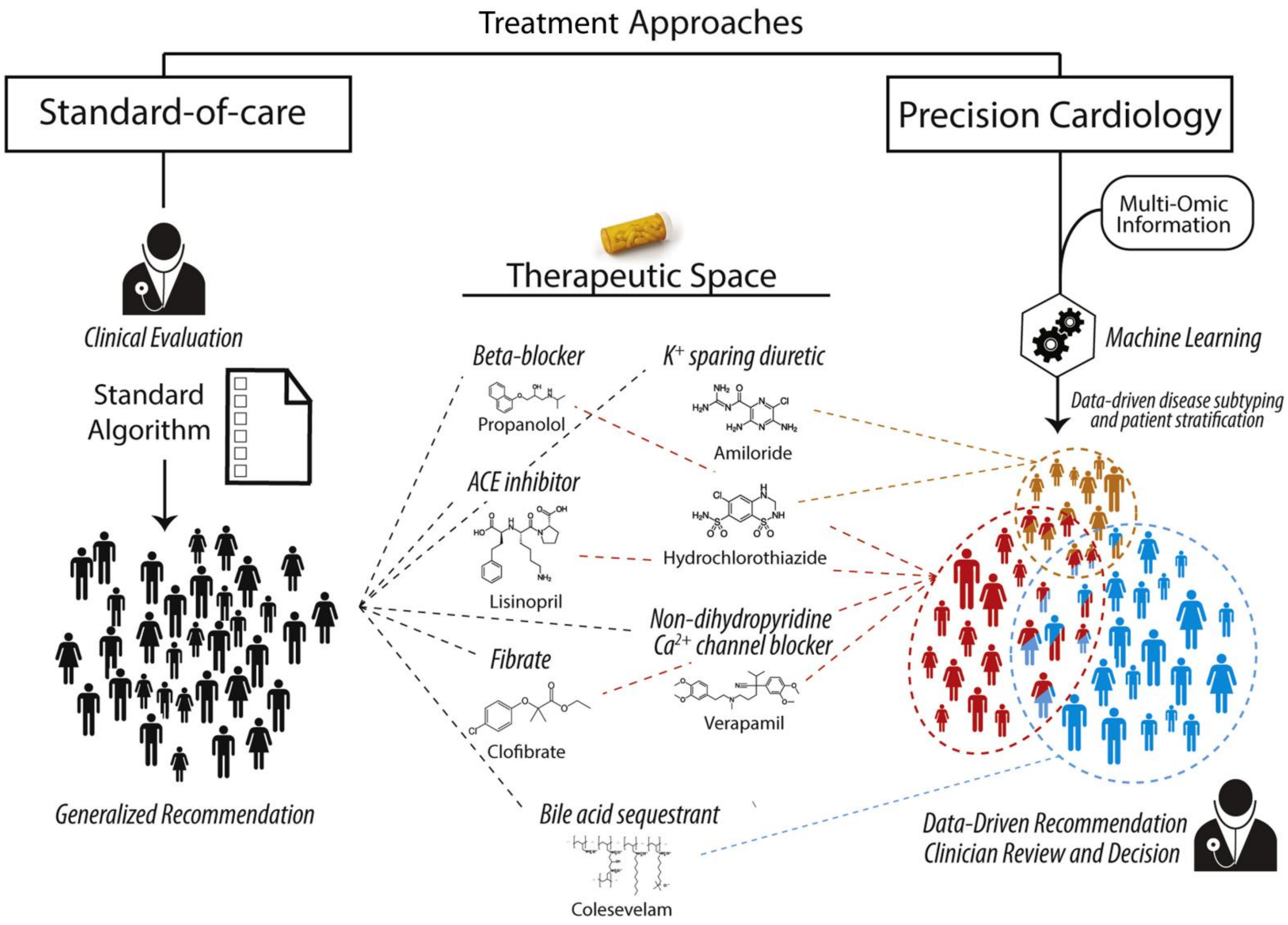
Assessing AI on clinical knowledge



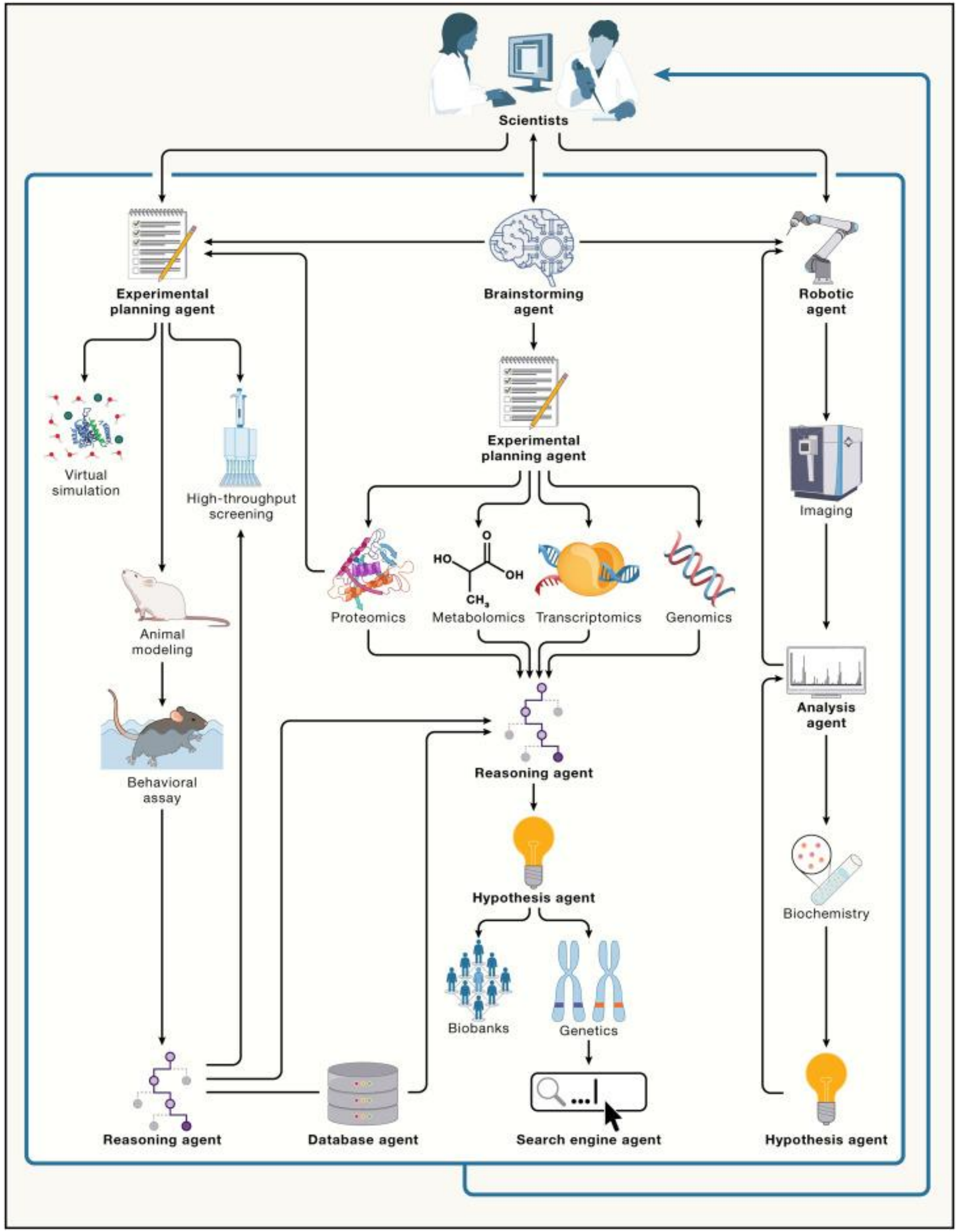
Large language models encode clinical knowledge, Nature 2023

General Vs Personalized Medicine

Enabling Precision Cardiology Through Multiscale Biology and Systems Medicine, JACC Basic Transl Sci,2017

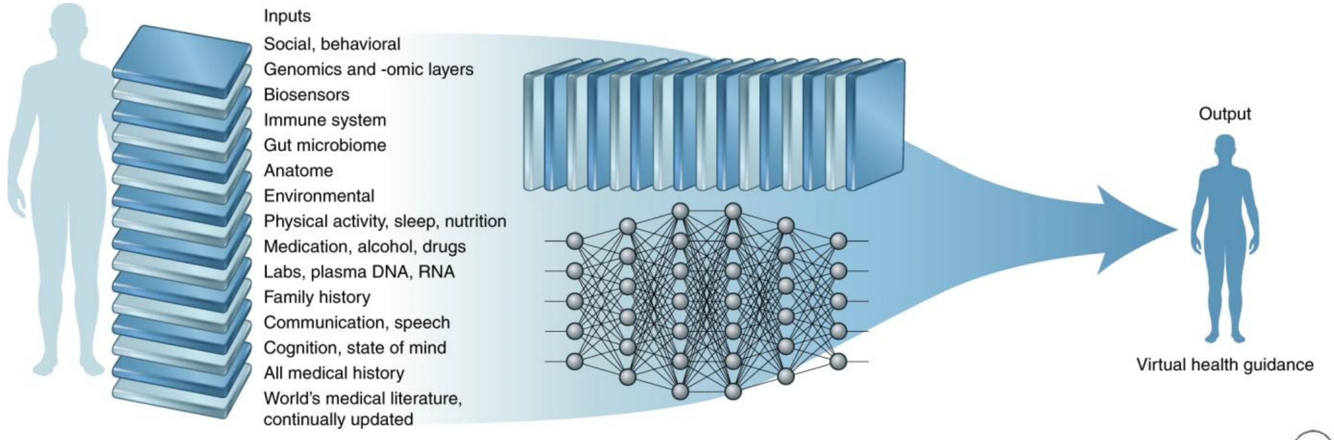


AI as an active agent of biomedical discovery



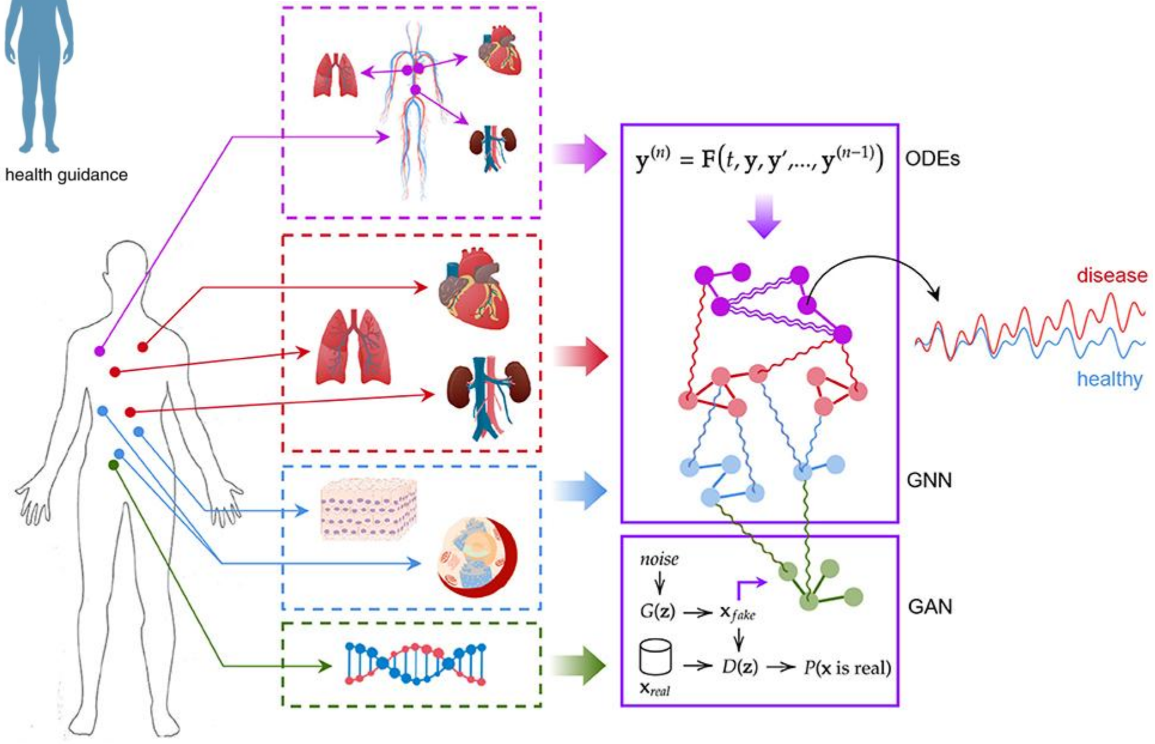
Empowering biomedical discovery with AI agents

Towards a multi-level and multi-systemic understanding of health



Topol, Nature Medicine 2019

Personalized digital twins modeling the human body as a whole and providing a panoramic view over individuals' conditions



Barbiero et al, Front. Genet.. 21

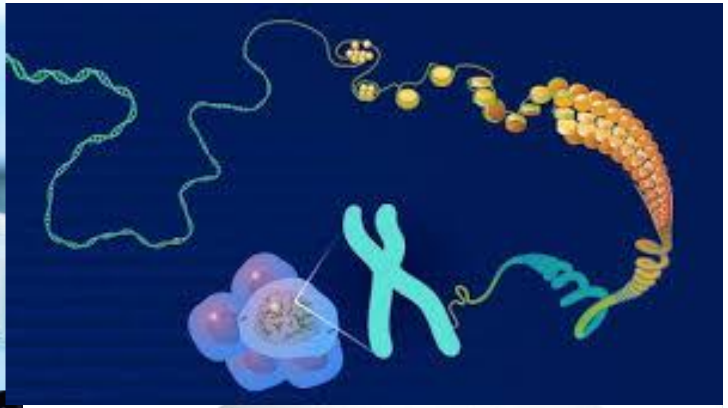
Perks of biomedical/health applications

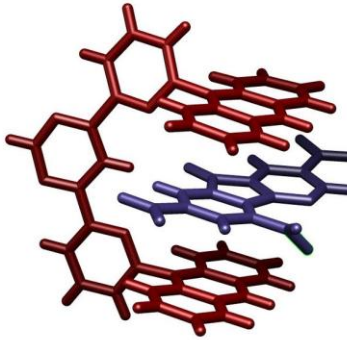
Biomedical data

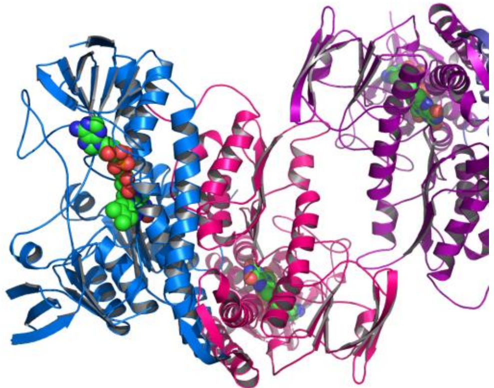
Complex

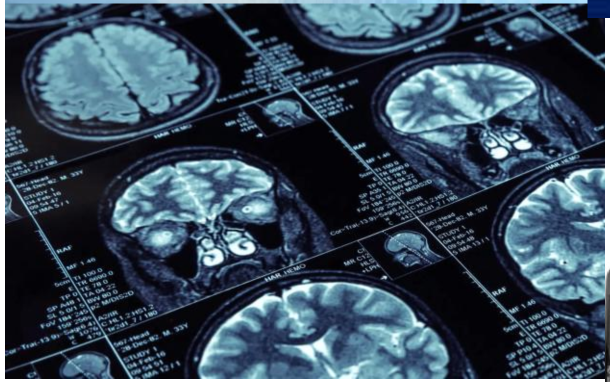
Multimodal

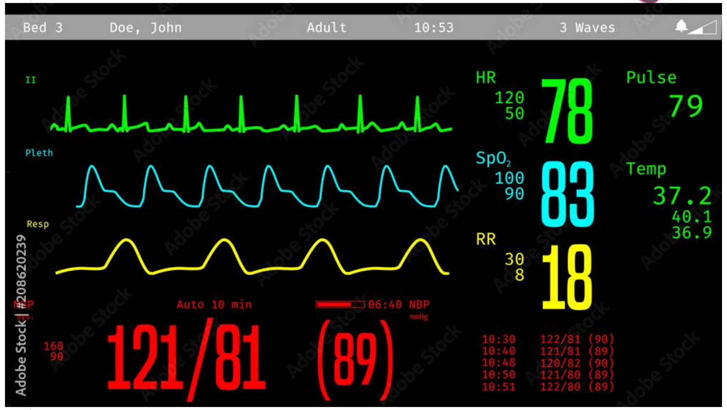
Heterogenous











Noisy

Biased and confounded

Long-tailed

What makes AI for health so different?

- Many questions are about unsupervised learning
  - Discovering disease subtypes, or answering question such as “characterize the types of people that are highly likely to be readmitted to the hospital”?
- Many of the questions we want to answer are causal
  - Naive use of supervised machine learning is insufficient
  - Performative effects
  - Assessing the impact of interventions

What makes AI for health so different?

- Very little labeled data
  - Recent breakthroughs in AI depended on lots of labeled data!
  - Motivates semi-supervised and self-supervised learning
- Sometimes small numbers of samples (e.g., a rare disease)
  - Learn as much as possible from other data (e.g., from healthy patients)
  - Model the problem carefully
- Lots of missing data, varying time intervals, censored labels

Trust, ethics and legal aspects

- Life or death decisions
  - Accuracy alone is no longer enough
  - Need robust, fair and accountable methodologies
  - Trustworthy AI (focus of other courses)
- Difficulty of de-identifying data
  - Need for data sharing agreements and sensitivity
- Difficulty of correcting for biases and inequities
  - Consideration of ethical and legal issues
  - Health data on which algorithms are trained are likely to be influenced by many facets of social inequality
- Rigorous testing and validation are needed
  - Longer time-to-market of technologies

Operational aspects

Difficulty of deployment of data-driven solutions

- Commercial electronic health record software is difficult to modify
- Data are often in siloes; everyone recognizes need for interoperability, but slow progress

Course execution and examination

Key ingredients of this course

- Weekly lectures (2/3) – D. Bacciu
- Weekly lab tutorials (1/3) – M. Podda
- Weekly lab exercises: home-work, voluntary, not-graded
- Monthly laboratory assignment: in-class, graded
- Monthly short methodological quizzes: in-class, graded

Lectures

- Introduction of concepts, terminology, models and methodologies
- We will use mathematical formalisms to introduce concepts, models and their inner workings (yes this will be part of the oral exam)
- We will not indulge in mathematical demonstrations but favour understanding of the methodologies and their use

Lab Tutorials

- Revisit and expand key concepts introduced in lectures with focus on their practical implementation and usage
- Mostly in Python with Jupyter notebooks (Colab preferably, but feel free to run your own notebooks locally)
- Prior knowledge of numpy and pandas is assumed from previous courses, but will be refreshed as we go by
- Will also leave some additional notebooks for home study
  - Optional, but still useful to complement the lectures

Weekly Lab Exercises

- Quick exercises (10 minutes approx.) in-class
- 1-2 per lab lesson
- Designed to get your hands dirty quickly
- You can work them out in groups if you wish
- Informal, not graded
- We'll discuss your (and my) solution at the end
- Additional notebooks will also contain exercises that you can work out at home

Laboratory Assignments

- More involved exercises (30 min to 1 hour) in-class
- 1 every 4 lab lessons
- Designed to test your understanding
- Work on them on your own
- Formal and graded, 0-3 points (see in a minute)
- Will publish a solution and grades in the following days

Methodology Quizzes

- Monthly quizzes based on the content of the lectures of the period
- Will be served in-class and they will be closed-book
- Will be short, requiring maximum 10 minutes to be completed and performed “on paper” (electronic devices not allowed)
- Will be graded on completion: each fully completed quiz scores 1 point (fractions of the point are possible)

They will be **delivered at random lectures** and unannounced

Course Assessment (Exams)

Preferential way for full time students

- Laboratory assignments - A total of 4 laboratory assignments related to lab hands-on
- Methodological quizzes - A total of 4 laboratory assignments related to lecture content
- Oral exam – An examination on the course program

Alternative way (part-time students, those who fail or don’t like the other way)

- Final Project – A coding project on a topic of interest for the course
- Oral Exam - A 15 minutes presentation of the final project plus examination on the course program

Final Project

- A coding project on a topic of interest for the course (defined by the instructors)
- What to deliver
  - The code (not the data)
  - A 10 pages report describing the project methodology and its validation
  - A 10/15 slides presentation (that will be given on the oral day)
- Timeline
  - Final project topic published: early May
  - Project materials delivered on Moodle by the exam date published on the esami platform (strict)

Oral Exam

- Give your presentation on the final project (15 minutes)
  - Discuss it in front of instructors and anybody interested
  - Questions on the methodology, validation and implementation

Only for those who did not choose the preferential way

- An oral exam with questions covering the course contents
  - Questions on models, methodologies and their implementation
  - Lectures whose content is not relevant for the final exam will be clearly marked as such

Grading – The Preferential Way

- Laboratory assignments grade G_{lab}
  - Each assignment is scored in [0,3]
  - Total grade achievable $G_{lab} = 12$ (since there are 4)
- Methodological quizzes grade G_{quiz}
  - Each quiz is scored in [0,1]
  - Total grade achievable $G_{quiz} = 4$ (since there are 4)
- Students are admitted to the oral if G_{lab}+ G_{quiz} > 8
- Oral exam grade G_{oral} is scored in [0,18]
- The final grade is $\min(G_{lab} + G_{quiz} + G_{oral}, 30\ Lode)$

Grading – The Alternative Way

- Final project grade G_{prj} is scored in [0,30]
- Oral exam grade G_{oral} is scored in [0,32]
- The exam is passed if both G_{prj} and G_{oral} are at least 18
- The final grade is

$$\frac{G_{prj} + G_{oral}}{2}$$

Policies

- You can collaborate with others, but we ask that you write your solutions individually
- We favor in-person attendance to all classes: preferential exam modality will support this
- We support use of genAI (LLMs, coding copilots, etc):
  - **Responsibility for content**: Students who use genAI in their assignments take full responsibility for the content they submit
  - **Acknowledgment of AI use**: Clearly acknowledge any use of genAI, specifying the nature and extent of assistance received from AI.
  - **Ethical use and originality**: Do not use AI to plagiarize, misrepresent original work, or fabricate data
  - **Instructor discretion**: We may specify assignments genAI use is prohibited

Next Lecture

Fundamentals of probability and statistics for AI

- Refresher on Probability and Statistics
- Hypothesis Testing and Statistical Inference
- Statistical dependence
- Statistics for healthcare data