George Khushf
University of South Carolina
Khushf@sc.edu

Anchoring the Ethical Debate About AI in Medicine

Two ways of approaching ethics of AI in medicine

I. Consider the discontinuities:

Step 1: Ask what's new about AI, e.g. Big Data and ML

Step 2: Understand and proactively mitigate adverse consequences of novel, potentially disruptive developments on medicine

II. Consider the continuities:

Step 1: Provide a historical framework for aligning contemporary and earlier visions for AI and medicine

Step 2: Situate recent developments associated with AI in relation to broader developments so ethical insights worked out for earlier stages can inform contemporary ethical reflection on AI and medicine

Outline:

I. A clinical case exhibiting the promise of AI Assistants in Medicine

II. A contemporary proposal for “Deep Medicine” - Eric Topol’s vision for how AI Assistants enable recovery of a caring relation between physicians and patients

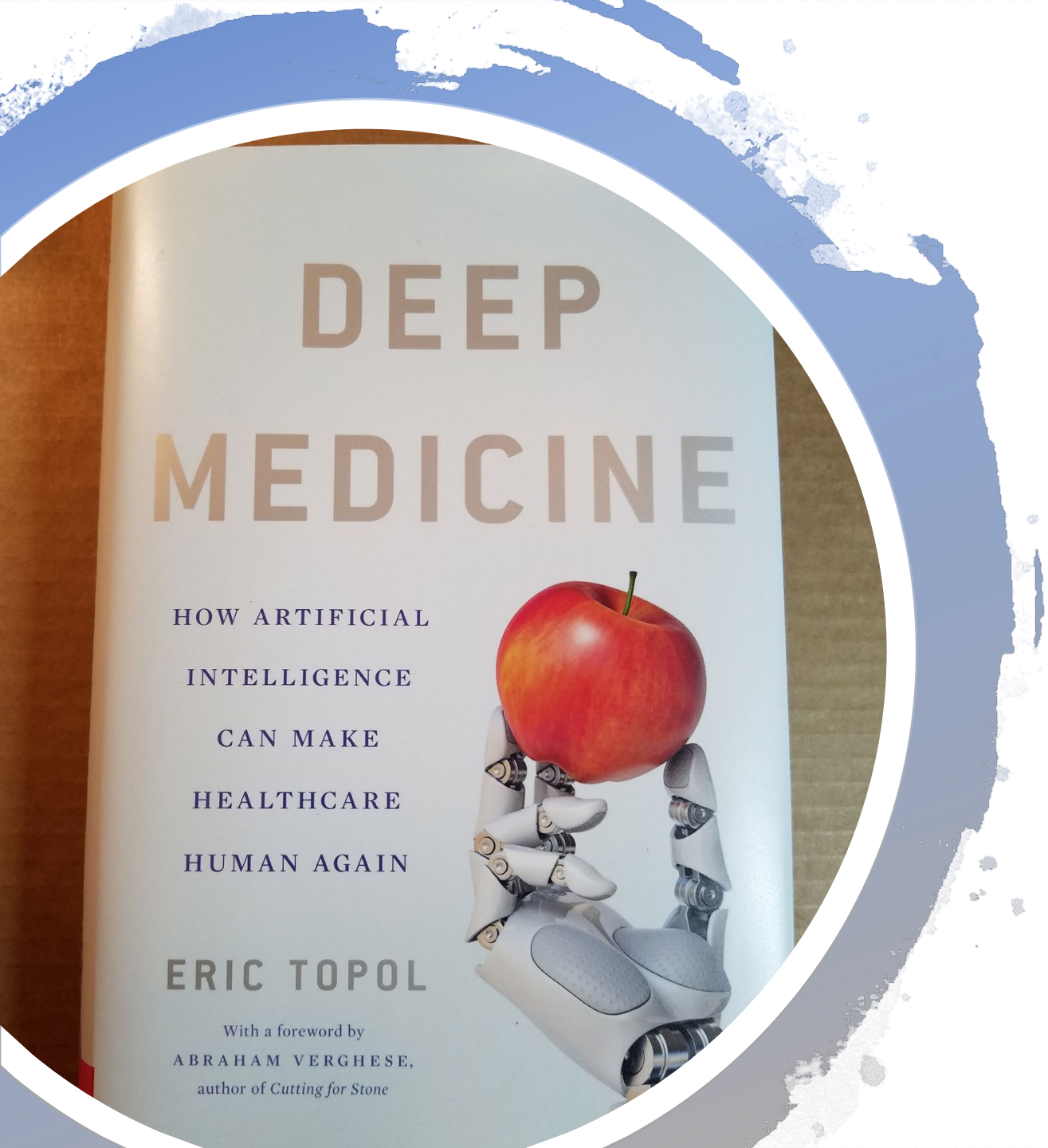
III. The encounter between AI and Medicine 1960-2010: A Historical context for AI assistants in medicine

IV. The ethics of AI in the 1980s

V. Lessons from earlier reflections for the ethics of AI

I. A clinical case in genomic medicine:

- 8 days after birth a boy presents with seizures at Rady Children's Hospital in San Diego, CA in USA
- Extensive efforts to diagnose and treat are unsuccessful; seizures are worse
- Stephen Kingsmore, a medical geneticist, sequences the whole genome of the boy
- 125 gigabytes of data are analyzed by AI to identify 1000 rare variants associated with disease
- Natural language processing AI identifies 88 phenotypic traits from Electronic Health Record (EHR)
- AI integrates genetic and EHR data to provide diagnosis: ALDH7A1 causes metabolic deficiency that leads to seizures. Treat with B6 and two amino acids, arginine and lysine.



II. Eric Topol's proposal for how AI Assistants will enable recovery of a caring physician-patient relation

Deep Medicine

Deep Phenotyping

Deep Learning

Deep Empathy

Why there has been “a steady degradation of the human side of medicine”

1. Loss of time
2. Loss of capacity to empathize with those who suffer
3. Loss of presence: genuine listening and attention to the stories of patients
4. Loss of rituals associated with physical exam
5. Loss of sustained physician-patient relation

Physicians are overwhelmed by demands of “shallow medicine” and by agendas of business and administrators of health systems which redirect their attention. Crucial is recovery of time!

AI enables “deep empathy”

AI functions as a “Medical Assistant” who will handle (better) the technical aspects of clinical reasoning

The human physician gets the time to interact with the patient and provide care that is responsive to patient suffering

Additional Concerns:

"Deep Liabilities"

- Explainability and the black box
- Privacy and security of data
- Bias and fairness Issues
- Transparent, contestable, accountable

III. AI and Medicine: 1960-1970s

- AI is oriented toward a practical goal: making agents that exhibit “intelligence” in the performance of a meaningful task.

- AI arose from a confluence of fields:
  - computers, information systems, interfaces
  - mathematics, logic and statistics
  - operations research
  - decision theory
  - cognitive science
  - Linguistics

- Medicine was one of the earliest fields to engage AI (broadly conceived): efforts were oriented toward designing Medical Diagnostic Support Systems (MDSS).

Earlier visions associated with AI in medicine:

- RS Ledley and LB Lusted:
  - "Reasoning foundations of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason," *Science* 1959
  - "The use of electronic computers in medical data processing: aids in diagnosis, current information retrieval, and medical record keeping," *IRE transactions on medical electronics* 1960
  - Society for Medical Decision Making (SMDM) founded in 1979, with its journal, *Medical Decision Making*. LB Lusted is first editor.
- Homer Warner:
  - *Computer-assisted Medical Decision-Making* (Academic Press, 1979).
  - Serves as CIO for University of Utah Health Sciences Center (later Intermountain Health System); site of first EHR system.
  - Launches first university bioinformatics program and organizer of the American College of Medical Informatics
  - Provides guidance for NIH programs advancing AI in medicine.

Reasons for developing MDSS

1. ***Advance quality and accuracy of clinical diagnosis***: "approaches that are systematic, complete, and able to integrate data from diverse sources"
2. ***Avoid error and make clinical decisions more reliable***: "making the criteria for decisions explicit, and hence reproducible"
3. ***Efficiency***: "balancing the expenses of time, inconvenience, or funds against the benefits and risks of definitive actions"
4. "Improve our ***understanding of the structure of medical knowledge***…"
5. "Improve our ***understanding of clinical decision making***, in order to improve medical teaching and to make computer programs more effective and easier to understand."

From Shortliffe, Buchanan, and Feigenbaum, "Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids" (1979)

Steps needed for MDSS

1. The “art” of clinical reasoning had to be reconstructed in an algorithmic form, providing step-by-step pathways leading to a diagnosis and recommended treatment.
2. The knowledge base of medicine had to be made explicit and encoded in information systems; this includes a large store of possible diseases, with linkage to the underlying basic sciences.
3. Medical records had to be digitized so they could be made accessible to a computer and easily shared.
4. Computers needed to be made broadly available
5. An interface between the human, computer, and information systems needed to be developed so a clinician could easily search for information, input relevant data on patients, and get answers from a diagnostic system.

The Hope

"One may hope that the computer, well-equipped to store a large volume of medical information and ingeniously programmed to assist in decision-making, will help free the physician to concentrate on the application of bedside skills, the management of the emotional aspects of disease, and the exercise of good judgment in the nonquantifiable aspects of clinical care."

From G. Anthony Gorry, "Computer-Assisted Clinical Decision Making", Methods of Information in Medicine 12: 45-51 (1973)

III. AI and Medicine: 1980-2010s

- The "Expert Systems" stage in the history of AI
- MDSS perform as well as specialists in specialty areas
- Example: the Stanford MYCIN Experiments
- MYCIN provided diagnosis of bacterial infections and recommendations for treatment.
- Reason for MYCIN:
  - primary care physicians often needed consults in this area
  - physicians frequently made errors when selecting antimicrobial agents.

Conclusions of the MYCIN Experiment

From: Bruce Buchanan and Edward Shortliffe, eds., *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* (Addison-Wesley1984).

According to criteria for assessing **performance**:

MYCIN was a **success**. It even passed a Turing Test: those seeking a consult on "hard cases had their cases randomly allocated to either MYCIN or infectious disease faculty at Stanford, and recipients found recommendations as appropriate as those from faculty. When outside experts evaluated the consults, they "disagreed with the MYCIN's recommendation no more than they disagreed with the recommendations from Stanford faculty."

According to criteria for **acceptability**:

MYCIN was a **failure**. There were practical reasons for failure related to availability of computers and ease of use. But the biggest reason for failure was MYCIN's **inability to provide adequate explanations.**

The broader indirect impact of the AI encounter on medicine:

- Physicians came to understand their own reasoning in terms of algorithms, and this sets new standards for a good explanation:
  NEJM editor, Jerome Kassirer and Richard Kopelman, *Learning Clinical Reasoning* (1991).

- Promotion of guidelines and clinical pathways:
  David Eddy's series in JAMA; *Clinical Decision Making* (Jones and Bartlett, 1996)

- Systems initiatives for advancing quality in medicine
  To Err Is Human: Building a Safer Health System (2000) and the following IOM reports

- Evidence Based Medicine (EBM)
  David Sackett, "Bias in Analytic Research," in The Case-Control Study Consensus and Controversy (1979).

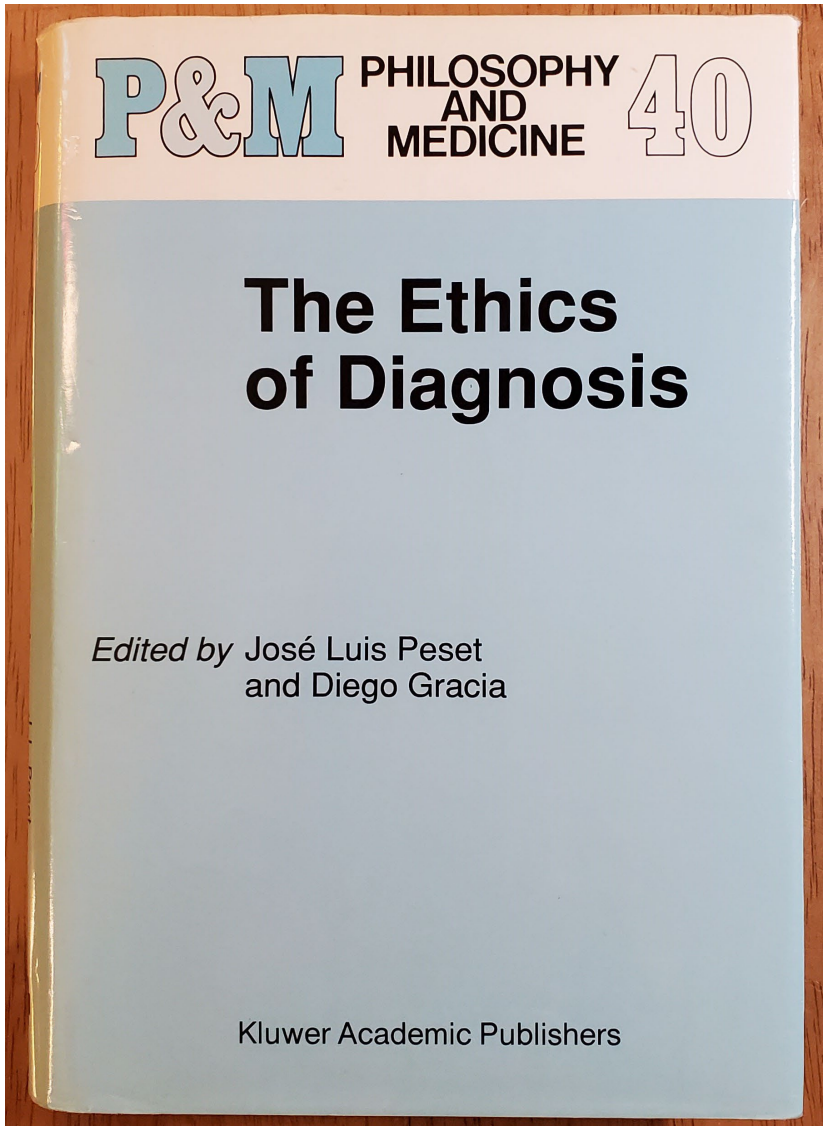
- Electronic Health Records (EHRs) and a computer between physician and patient

IV. The *ethics* of AI from the 1970s:

"The physician-patient relationship is fast being converted from a diadic to a triadic relationship in which the computer promises to be an active, efficient, helpful, and not so silent partner. To what extent will the physician be replaced by his mechanical partner? What realm of clinical decision-making remains uniquely the physician's." (p. 173)

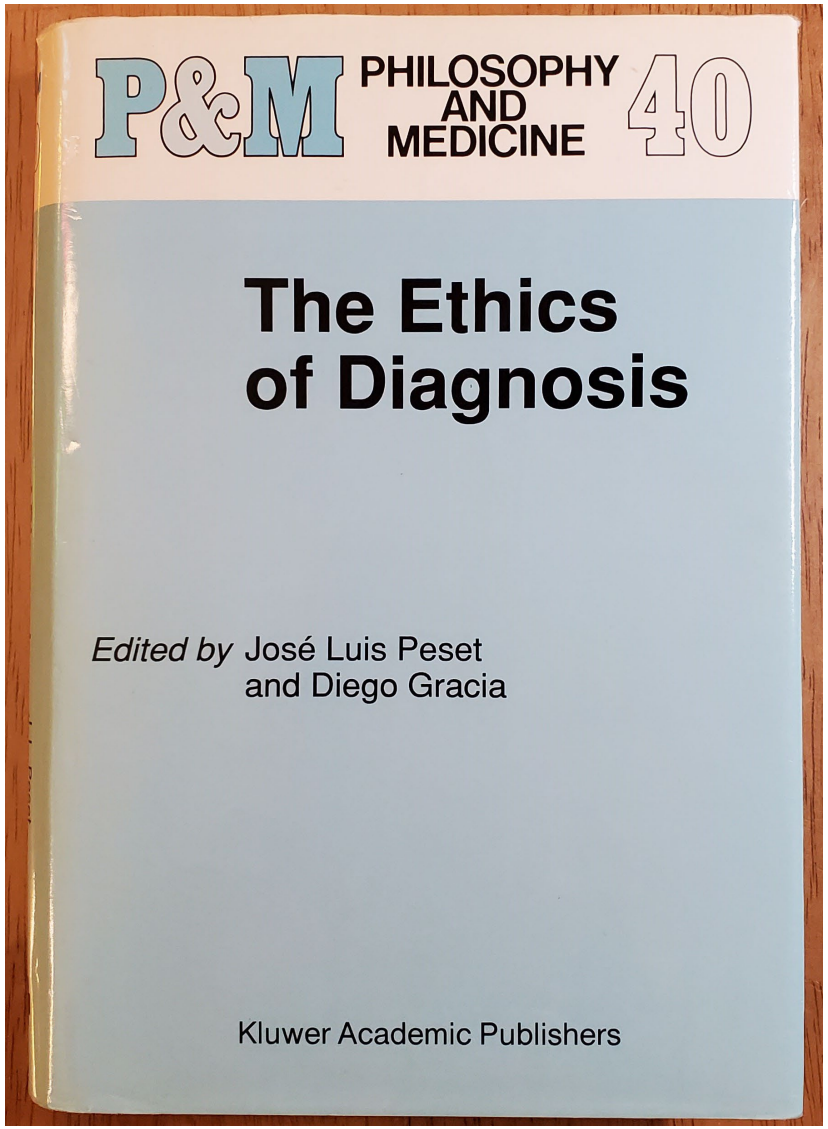
"We can assume that the computer will be a partner in the physician-patient relationship aiding and even replacing the physician partly or completely" (p. 175)

But: "The availability of computer diagnosis and decision-analysis need not be 'dehumanizing' especially if the patient preferences are included. Computers allow the clinician more time for the crucial third stage of clinical judgment – recommending and executing the right and good action for *this* patient." (p 189)



From Edmund Pellegrino, “Value Desiderata in Computer Diagnosis”, in Peset and Gracia (eds), The Ethics of Diagnosis (Kluwer, 1992).

This publication arose from a series of workshops on the ethics of computer diagnosis and AI from the 1980s. These, in turn, addressed proposals from the 1970s for how AI would transform medicine.



Insights on the ethics of AI from:

Mary Ann Gardell Cutter
H. Tristram Engelhardt
Drew Leder
Edmond Murphy
Edmund Pellegrino
Kenneth Schaffner
Stuart Spicker
Henrik Wulff

“Patient preferences must modify the conduct and conclusion of any diagnostic schema, computerized or not. Suffice it to say that every clinical decision involves an intersection of value systems – of the physician, patient, nurse, family, friends, colleagues, and society.”

Edmund Pellegrino, “Value Desiderata in Computer Diagnosis” (p. 183)

"These modulations of logical formality by the habits of human thinkers are operative in all decision-making. In clinical decisions they superimpose themselves like a grid over all the intricate steps in the process of clinical judgment. The key to the intrinsicality of clinical decisions is their telos – a technically correct and morally good healing decision for *this* patient, at *this* time, and in *this* context."

Edmund Pellegrino, "Value Desiderata in Computer Diagnosis" (p. 186)

"[T]he signs, symptoms, and tests the clinician chooses to enter into the Bayesian calculus are determined by the instrumental end they serve. Not only are the data not independent pathophysiologically, but they are linked in a value context projected backward from the goal of doing what is good for *this* patient."

Edmund Pellegrino, "Value Desiderata in Computer Diagnosis" (p. 187)

V. Our ethical challenge today:

… to situate the discontinuities – the novel capacities, opportunities, and challenges – in relation to the continuities, so we can best understand, utilize and further refine the powerful AI tools that lie within our immediate horizon. The ideals of individualized humane care presented in The *Ethics of Diagnosis* can serve as a focal point for clarifying the values and structuring of clinical agency so Topol's ideal of humane care is realized by his AI assistant, rather than undermined by it.

Common themes over the decades:

1. Bias … in data sets, e.g. those used for machine learning; values encoded in both the data and in algorithms, but also those of agents and those informing individualized care for *this* patient
2. Privacy, security and surveillance of patient information
3. Promise for expanding equity, access and for improving quality
4. Question of whether AI should be surrogates for clinicians, assistants, or hybrid roles? (what is scope of practice, shifting professional roles)
5. Opacity of AI to agents using it, thus need for explainability and transparency … and why this is important for individualizing care
6. Managing expectations (hype vs skepticism) and developing a balanced response so promise of AI is realized without disrupting ethical fabric of the clinical encounter

We should also consider all of the unintended effects of AI:

1. Does EHR just enable and expand capacities of medicine? Do you have more time with patients for that humane part of care?
2. Problems raised by integration of clinical and economic considerations into systems developments and decision making
3. Odd effects of EBM on care (tensions between ideals of science and of individualized decision making that is responsive to patient values)
4. To what degree does Pellegrino's account of prudential reasoning still convey how physicians should reason? How they do reason?
5. Current AI (associated with Big Data and ML) potentially magnifies the irrational aspects that were regarded as the most problematic part of the earlier "expert systems" stage of AI.

Key problem with ML: it is not explainable in the algorithmic terms physicians internalized from their earlier encounter with AI.

Future directions for the ethics of AI

Commenting on work by Barbara Grosz that is oriented toward design of AI assistants, Lily Hu notes:

"'What we need to think about is how technologies embed particular values and assumptions. Exposing that is a first step: realizing that it's not the case that there are some ethical questions, some non-ethical questions, but really that, in everything we design ... there are always going to be normative questions at hand, every step of the way.' Integrating that awareness into [clinical practice] is critical to ensuring that 'the world that we're building, with ubiquitous technology, is a world that we want to live in. ' "

On the Embedded EthiCS Initiative, from *Harvard Magazine* (Jan-Feb, 2019), pp. 44-49, 77

A warning from Pellegrino:

The physician “need not worry about being replaced by the machine nor about its dehumanizing effects. Computer diagnosis and decision analysis may or may not require fewer people. But like every new technological advance in clinical medicine, they demand more of people as people, more of their humanity, not less. Computer assisted decisions place the highest premium on those things only humans can do, i.e., empathizing, educating, caring for, making decisions with as well as for patients.” (p. 190)

And a warning about the warning: whose “assistant”?