NEJM AI

NEJM AI 2024; 1 (3)
DOI: 10.1056/AIp2400036

PERSPECTIVE

To Do No Harm — and the Most Good — with AI in Health Care

Carey Beth Goldberg,[1] Laura Adams, R.N., M.S.,[2] David Blumenthal, M.D., M.P.P.,[3] Patricia Flatley Brennan, R.N., Ph.D.,[4,5] Noah Brown,[6] Atul J. Butte, M.D., Ph.D.,[7] Morgan Cheatham,[8] Dave deBronkart,[9,10] Jennifer Dixon, M.D., Ph.D.,[11] Jeff Drazen, M.D.,[12] Barbara J. Evans, Ph.D., J.D., L.L.M.,[13] Sara M. Hoffman,[6] Chris Holmes, Ph.D.,[14,15] Peter Lee, Ph.D.,[16] Arjun Kumar Manrai, Ph.D.,[6,12] Gilbert S. Omenn, M.D., Ph.D.,[17] Jonathan B. Perlin, M.D., Ph.D.,[18] Rachel Ramoni, D.M.D., Sc.D.,[19] Guillermo Sapiro, M.S., Ph.D.,[20,21] Rupa Sarkar, Ph.D.,[22] Harpreet Sood, M.B.B.S., M.P.H., M.R.C.G.P., F.F.C.I.,[23,24] Effy Vayena, Ph.D.,[25] Isaac S. Kohane, M.D., Ph.D.,[6] for the RAISE Consortium*

Received: January 10, 2024; Accepted: January 11, 2024; Published: February 22, 2024

Abstract

Drawing from real-life scenarios and insights shared at the RAISE (Responsible AI for Social and Ethical Healthcare) conference, we highlight the critical need for AI in health care (AIH) to primarily benefit patients and address current shortcomings in health care systems such as medical errors and access disparities. The conference, embodying a sense of responsibility and urgency, emphasized that AIH should enhance patient care, support health care professionals, and be accessible and safe for all. The discussions revolved around immediate actions for health care leaders, such as adopting AI to augment clinical practice, establishing transparent financial models, and guiding optimal AI use. The importance of AI as a complementary tool rather than as a replacement in health care, the necessity of responsible patient data usage, and the potential of AIH in improving access to care were stressed. We underscore the financial aspects of AIH, advocating for models that align with care improvement. Specific and practical next steps and decisions are provided for each major issue. We conclude with a call for ongoing dialogue and ethical commitment from all stakeholders in AIH, reflecting on AI's promise for health care advancement and the need for inclusivity and continuous evaluation in its implementation.

Introduction

A cancer patient plugs his medical information into a chatbot and learns that his oncologist has misidentified his tumor subtype and thus prescribed the wrong therapy.

A health system leader hears from early-adopter clinicians that using AI to record and summarize patient visits vastly improves the visits.

The author affiliations are listed at the end of the article.

Dr. Kohane can be contacted at isaac_kohane@harvard.edu or at Department of Bioinformatics, Harvard Medical School, 10 Shattuck St., Boston, MA 02115.

**The names and affiliations of the members of the RAISE Consortium are listed in the Appendix.*

A physician in Latin America launches a project that uses cellphone-based AI to support mental health among remote rural residents with no other access to care.

These real-life stories, recounted at a recent international conference on the use of AI in health care (AIH), underscore the prime directive to emerge from the gathering of more than 70 health care and technology leaders: AIH must first and foremost directly benefit patients, and it is an ethical imperative to make sure it does so. The goal is that benefits are maximized across all walks of society and harms are minimized, to the extent practical constraints will allow. Among the harms to be minimized are those inflicted by the current health care system, from medical errors to lack of access.

Given such ongoing problems and recent technological leaps, adopting AIH where it could help is a matter of urgency. At a moment when the complexity of modern medicine has surpassed the capacity of the human mind, *only* AIH will be able to perform many tasks. AIH thus seems to offer unparalleled potential for further medical progress, including for precision medicine — the right therapy, for the right patient, at the right time. Thus, it is in the interest of the public and the medical profession to hasten its adoption, so long as it is used safely and made maximally accessible for all. It is also urgent to determine how AIH can best deliver tangible benefits, including how it can help improve health and save lives in ways that will not otherwise happen. The public should not only be aware of this quest but also participate in it.

This consensus emerged at the RAISE (Responsible AI for Social and Ethical Healthcare) conference, which was organized by the Department of Biomedical Informatics at Harvard Medical School. The conference was held October 29–31, 2023, in Cape Neddick, Maine. The conference agenda and attendees can be found at https://dbmi.hms.harvard.edu/events/raise-symposium.

Multiple efforts are underway internationally to pass laws and set guidelines for the responsible use of the new AIH, including at the World Health Organization.[1] Among many other advisory bodies, the National Academy of Medicine (with the Coalition for Health AI) is currently developing a code of conduct that will harmonize efforts into one framework, with a special emphasis on ethics, equity, and vigilance for errors that may emerge once the AI is in use.[2]

The RAISE conference did not aim to reach a comprehensive consensus on all aspects of AIH. Participants focused instead on a handful of specific aspects of AI for health care delivery that are most relevant to the public. By incorporating diverse perspectives from the health care industry and beyond, the participants aimed to generate just a few broad principles and calls to action for public consumption, some of which can be applied immediately and others that can inform future, more granular guidance.[3] With the publication of the current article in two journals — *NEJM AI* and *Nature Medicine* — the goal was to stimulate an open discussion on these topics that can inform policymakers, companies, and professional societies in their own construction of policy and practice.

The Sense That "This Is the Future"

Infusing the spirit of the conference was a strong desire to avoid past mistakes in applying new technology and policies to medicine. For example, in the United States, electronic medical records have tended to serve billing needs more often than enhancing care. In addition, major privacy laws such as the European Union's General Data Protection regulation and the U.S. Health Insurance Portability and Accountability Act Privacy Rule are widely misunderstood and add administrative complexity that hinders AIH research.[4] Early AIH experiences showed that even powerful new technology will fail to catch on broadly and institutionally if health care staffers, payers, patients, or caregivers do not trust it or if they worry about liability.

Also widespread among conference attendees was a pronounced positive attitude toward the apparent potential of AIH, the sense that "this is the future." Some participants reported that this emphasis on benefits contrasted somewhat with a greater European regulatory emphasis on risks. That U.S.-predominant view accorded more with China, where more trials of AIH are registered than in any other country. Others pointed out that cellphone-based AI holds particular promise for people in both the developed and developing world who suffer from the worst health inequity: zero or poor access to quality care.

Underpinning it all was a sense of responsibility shared by clinical professions (medicine, nursing, pharmacy, and more) that are largely self-regulating. Ideally, according to broad consensus at the conference, professionals and patient advocates will determine best clinical practices and convey them to regulators, rather than vice versa. At the same time, even ideal professional norms can break down when misaligned payment systems promote harmful

overuse (in fee-for-service systems) or underuse, hence the critical importance of bringing the public into the discussion, to defend its own interests, and the appropriate oversight by external bodies.

To that end, the conference shares several key points.

1. WHO BENEFITS?

Conference members proposed that the central criterion for use of AIH should be the same as for other medical technologies — whether it will further four key aims: better health for all patients; better health care experience, including faster access; lower costs of delivering care and improved productivity; and support for the health care workforce, amid health care labor shortages and burnout.

Immediate calls to action for medical and health system leaders:

A. We recommend that health systems, health plans, and physician groups review and, if deemed appropriate, adopt AI to augment clinical practice. "Low-hanging fruit" includes enhancement of the doctor–patient interaction, including capture of the recorded patient visit, prioritizing and analysis of test and imaging results, differential diagnosis, plan for therapy and discussion of alternatives, instructions for the patient and caregivers, appointment scheduling and other administrative functions, and responses to patient questions at optimal levels of literacy.

B. Financial models for reimbursement for using AI should be transparent and evaluated annually. Identifying fully open and fair payment paradigms now is absolutely crucial.

C. The institutions should establish guides for clinicians, trainees, and patients on opportunities and optimal use of AI. Patients and other stakeholders should also push for widespread education and modeling of how to use AIH.

D. They should also create clear outcome expectations to provide ways to verify that the AI is indeed serving patients' and providers' interests and priorities, not only the fiscal pressures of private health care

(continued)

(continued)

systems or the budgetary constraints of a government-funded health care system. Publicly defined metrics should be assessed rigorously and shared across health systems.

2. SEPARATE VOICE OR FRIENDLY HELPER?

On the question of whether AI should be considered an additional entity in the traditional clinician–patient dyad, most participants agreed that the answer is "no." At this point, rather, AI should be seen as a tool supporting the clinician or an aid to patient self-management rather than as a new entity. AIH is complementing, not substituting, although that could change in the foreseeable future, subject to appropriate safeguards.

Calls to action:

A. **When providers use AIH:** Case law is still developing around the world, but at this point it seems that the clinician remains legally responsible for medical actions and decisions, even when AI is used in health care. However, if AIH is to be widely adopted, technology companies must accept some portion of the legal liability if an AIH system leads to harm and it is at fault. Also, the next generation of clinicians must be trained on how best to use AI to supplement their medical expertise.

B. **When patients use AIH:** Clarify that, similarly, the technology companies must move toward accepting some responsibility for outcomes when patients use the AI directly, just as with any other direct-to-consumer health tool or resource.

3. ABOUT PATIENT DATA

Because it is so important to train AIH models with broad and diverse data, AI models will likely rely extensively on data generated during the course of clinical care. When it comes to arrangements for ingesting patient data into AI models, conference participants strongly supported the use of "opt-out" — that is, the default is for patient data to be included — rather than "opt-in," which could lead to

further disparities in data collection because of historical mistrust of some populations toward medical science. Participants do, however, recognize that opt-out consent cannot fully eliminate the potential for harmful bias because populations harboring this mistrust may disproportionately exercise their right to opt-out, leaving them underrepresented. Whether opt-in or opt-out, consent cannot substitute for appropriate technical safeguards for privacy and responsible data use practices that will make AIH tools trustworthy and reduce the public's urge to opt-out.

There was also a broad consensus that builders of medical databases have long left out too many members of underrepresented minorities. Strong communities are now working on that issue, and hope is high that such past wrongs can be righted. A broad communication and education strategy is needed to explain how AIH helps "people like you," and to involve underrepresented groups in proactively building databases.

Calls to action:

A. Lean toward opt-out models (already the industry standard) but be sure patients are well-informed about their options, including the right to say "No, you may not use my data in training your model."

B. Determine how to measure (or at least estimate) and address the privacy risks that emerge when patient data are used to train models. Guarantees from AI vendors are needed that data will be protected, patients will not be identified, and there will be deterrent penalties for data misuse.

C. Review and consider using guidelines that have already been developed for proper ethical use of patient data; for example, U.K. health authorities have developed a program for ensuring that data represent population diversity[5] and another called "The Five Safes"[6] for using public data for the public good.

D. AIH performance critically depends on the diversity of the data and the populations on which they are based. As AI models grow, they will need dramatically more multimodal training data from everywhere, and they will need to be transparent about their

(continued)

(continued)

performance for different demographic characteristics. Thus, their developers need a well-regulated data economy that tracks provenance and benefits across society, ensuring that all are included in the data pipeline, even hard-to-enlist populations such as the uninsured or migrants in the European Union.[7] The data economy developers must always think about who is missing and report on that aspect of the model development to the public.

4. DIRECT CONSUMER ACCESS

The consensus of conference participants was that AIH supports but does not currently replace clinician-provided care or patient self-management. The ubiquity of smartphones and Internet connections brings access to AIH tools to almost everyone. Where care access is limited by resources, including in wealthy countries with staff shortages, AI is a valuable alternative — if handled correctly. Patients can now enjoy round-the-clock access to extensive medical information and insights far more sophisticated than online health information platforms such as WebMD. This is a sea change in a field in which, traditionally, patients have not only faced barriers in accessing their own medical data but also lacked inexpensive means to reason with those data.

Calls to action:

A. Require that AI vendors reveal the sources (as in a nutritional label) of their training data and the overarching instructions guiding the AI, in ways accessible to the patients themselves and also to regulators, to ensure that it is representative and the alignment is explicit.

B. Bring patients, patient advocates, and clinicians into the process of designing the technology at every stage, particularly in setting its guardrails and limits on its use, both to ensure that it suits their needs and to build trust.

C. Clinicians should now expect that many patients will arrive with information gleaned from AI. They should encourage the common practice of patients checking the online visit

(continued)

(continued)

summary of diagnoses and planned therapies after the visit. (It would be wise occasionally to check what the patient is finding!)

D. Remember that it may be advantageous for patients to avoid a visit to a provider by using an AI-powered alternative, shifting some time-held practices, and disrupting the financial models underlying health care. Health systems should strive not just to improve provider care with AI but eventually to substitute for some of it. A participant noted that many in their country would actively like to avoid doctors — a widespread sentiment elsewhere as well.

5. MONEY MATTERS

AIH incurs costs, including for its development, data acquisition, and the ongoing evaluation and monitoring of its quality and impact. Thus far, there is no single standard by which payments are made. Some models for the acquisition of AIH tools include outright purchase, pay-per-use, and subscription services. Even where health care is fully subsidized by the government, AIH tools might be funded or supplied by not-for-profit or for-profit companies.

Experience shows that payment schemes can influence real-world uptake of medical innovations. Earlier AI tools often met with limited demand, particularly if they were pay-per-use, leading to underuse; there is also a risk, however, that payment models can provide incentives that lead to harmful overuse, either of the AI tool or products recommended by it, regardless of whether the government pays directly for health care.

Calls to action:

We recommend considering:

- Favoring subscription or up-front payment models rather than bill per use. As Harvard University health economist David Cutler told the conference, "Do not fight the monetary incentives. You will lose."

(continued)

(continued)

- Reimbursement should be tied to problems solved and care improvements: Third parties (e.g., academics, nonprofits) will have to erect the infrastructure (e.g., journals, reports) to track over the coming years whether the new AI actually leads to reduced administrative costs, improved patient outcomes, improved therapeutic options, and more. The infrastructure should also include postmarketing surveillance by regulators and structured evaluations funded by public research and development bodies, as well as more international learning on AIH developments.

Finally, speaking more generally, if AI's potential for improving health turns out to be as great as it now seems, then arguably all health care leaders bear the responsibility to support ways to help patients and staff take advantage of it.

It is also an ethical imperative for all involved with AIH to focus on the public good. For example, drugmakers should provide medications at free or low cost to poor countries. Similarly, AIH companies should make an industry-wide commitment to develop applications and provide them for free or low cost to underserved populations that cannot otherwise afford them.

Given the growing complexity of health care, AI seems to provide the most promise for its continued improvement. The technology is advancing at a blistering speed and, considering the tremendous uncertainty about how it will develop in the coming years, there will be an ongoing need for conversations on topics such as those explored here. Conclusions reached today may no longer be relevant even 1 year from now. Lives may be saved, however, that would not have been saved even just last year. So, while the conversation around how best to use AIH for public and patient good needs to continue, it also needs to begin now.

Disclosures

Author disclosures are available at ai.nejm.org.

Appendix

The members of the RAISE Consortium are as follows: Emily Alsentzer, Ph.D., Brian Anderson, M.D., Ran D. Balicer, M.D., Ph.D., M.P.H., Andrew

L. Beam, Ph.D., Erwin Bottinger, M.D., Rebecca W. Brendel, M.D., J.D., Payal Chandak, Arnaub Chatterjee, M.H.A., M.P.A., Michael Chernew, Ph.D., Antoine Duclos, M.D., Ph.D., Lee A. Fleisher, M.D., William Gordon, M.D., Elizabeth Healey, Alejandro Hernández-Arango, M.D., M.S., Michele Kathleen Herndon, M.S.N., R.N., Alexander Hoffmann, M.B.A., Edward M. Hundert, M.D., Indra Joshi, M.B.B.S., Tobias Kowatsch, Ph.D., M.S., Stephen Kraus, M.B.A., Lisa Soleymani Lehmann, M.D., Ph.D., M.Sc., Karim Lekadir, Ph.D., Vincent X. Liu, M.D., M.S., Daniel J. Nigrin, M.D., M.S., Reut Ohana, M.B.A., M.Sc., Nigam H. Shah, M.B.B.S., Ph.D., Haris Shuaib, M.Phys., M.Sc., Tania Simoncelli, M.S., Amelia Li Min Tan, Ph.D., Donrich Thaldar, Ph.D., Eugene Tunik, Ph.D., P.T., Tommy Wang, Ph.D., John Wilbanks, Yuchen Xu, and Jianfei Zhao, Ph.D.

The affiliations of the members of the RAISE Consortium are as follows: Brigham and Women's Hospital and Harvard Medical School (E.A.); Chief Digital Health Physician, MITRE Corporation and Co-Founder, Coalition for Health AI (B.A.); Clalit and Berkowitz Living Lab (R.D.B.); Department of Epidemiology, Harvard T.H. Chan School of Public Health (A.L.B.); Wyss Center for Bio and Neuroengineering (E.B.); Harvard Medical School Center for Bioethics (R.W.B.); Harvard-MIT Program in Health Sciences and Technology (P.C.); TriNetX (A.C.); Harvard Medical School (M.C.); Research on Healthcare Performance Lab INSERM U1290 Lyon 1 University, Health Data Department Lyon University Hopital, France/Center for Surgery and Public Health, Brigham and Women's Hospital, Boston (A.D.); Emeritus Professor of Anesthesiology and Critical Care, Perelman School of Medicine of the University of Pennsylvania, Visiting Fellow, Duke Margolis Center for Health Policy, Duke University (L.A.F.); Department of Medicine, Brigham and Women's Hospital and Department of Biomedical Informatics, Harvard Medical School (W.G.); Harvard-MIT Health Sciences and Technology Program (E.H.); University of Antioquia, Alma Máter Hospital, Medellín, Colombia (A.H.-A.); Undiagnosed Diseases Network Foundation (M.K.H.); HMS (A.H.); Harvard Medical School Center for Bioethics (E.M.H.); BMJ Leader, Palantir (I.J.); (1) Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland, (2) School of Medicine, University of St. Gallen, St. Gallen, Switzerland, and (3) Centre for Digital Health Interventions, Department of Management, Technology, and Economics at ETH Zurich, Zurich, Switzerland (T.K.); Bessemer Venture Partners (S.K.); Brigham and Women's Hospital, Harvard Medical School, Harvard T.H. Chan School of Public Health, Boston (L.S.L.); (1) Universitat de Barcelona, Artificial Intelligence in Medicine Lab (BCN-AIM), Department of Mathematics and Computer Science, Barcelona, Spain, and (2) Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona, Spain (K.L.); Kaiser Permanente Division of Research (V.X.L.); MaineHealth, Tufts School of Medicine (D.J.N.); Clalit Research Institute, Innovation Division, Clalit Health Services, Tel Aviv (R.O.); Stanford Health Care, Palo Alto, CA (N.H.S.); Guy's & St Thomas' NHS Foundation Trust (H.S.); Chan Zuckerberg Initiative (T.S.); Harvard Medical School, Department of Biomedical Informatics (A.L.M.T.); School of Law, University of KwaZulu-Natal, Durban, South Africa (D.T.); Institute for Experiential AI and Bouve College of Health Sciences, Northeastern University (E.T.); Gordon and Betty Moore Foundation (T.W.); Milken Institute (J.W.); Shanghai Digital Medicine Innovation Center, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine (Y.X.); and *NEJM AI*, Boston and Jiahui Medical Research and Education, Shanghai (J.Z.).

The RAISE symposium was supported by grants from the Harvard Medical School Center for Bioethics, the Harvard Medical School Office of the Dean, the Gordon and Betty Moore Foundation, The Health Foundation, Microsoft Corporation, and Apple.

Author Affiliations

[1] Massachusetts Institute of Technology, Cambridge
[2] National Academy of Medicine, Washington, DC
[3] Department of Health Policy and Management, Harvard University T.H. Chan School of Public Health, Boston
[4] National Library of Medicine, Bethesda, MD
[5] University of Wisconsin–Madison, Madison
[6] Department of Biomedical Informatics, Harvard Medical School, Boston
[7] Bakar Computational Health Sciences Institute, University of California, San Francisco, San Francisco
[8] Warren Alpert Medical School of Brown University, Providence, RI
[9] e-Patient Dave, LLC
[10] Society for Participatory Medicine, Pembroke, MA
[11] The Health Foundation, London
[12] NEJM Group, Waltham, MA
[13] Levin College of Law and Wertheim College of Engineering, University of Florida, Gainesville
[14] Department of Statistics and Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[15] The Alan Turing Institute, London
[16] Microsoft Corporation, Redmond, WA
[17] University of Michigan Health System, University of Michigan, Ann Arbor, MI
[18] The Joint Commission, Oakbrook Terrace, IL
[19] U.S. Department of Veterans Affairs, Washington, DC
[20] Duke University, Durham, NC
[21] Apple, New York
[22] The Lancet Ltd., Lancet Digital Health, London
[23] National Health Service England, Hurley Group, Redditch, United Kingdom
[24] Huma, London
[25] ETH Zurich, Zurich, Switzerland

References

1. World Health Organization. Regulatory considerations on artificial intelligence for health. 2023 (https://iris.who.int/handle/10665/373421).
2. Coalition for Health AI. Blueprint for trustworthy AI implementation guidance and assurance for healthcare. December 7, 2022 (https://www.coalitionforhealthai.org/papers/Blueprint%20for%20Trustworthy%20AI.pdf).
3. GOV.UK. Policy paper: the Bletchley Declaration by countries attending the AI Safety Summit, 1–2 November 2023. November 1, 2023 (https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023).
4. The Commonwealth Fund. The electronic health record problem. December 13, 2018 (https://www.commonwealthfund.org/blog/2018/electronic-health-record-problem).

5. STANDING Together Working Group. STANDING Together. 2023 (https://www.datadiversity.org).
6. UK Data Service. What is the Five Safes framework? (https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework).
7. Publications Office of the European Union. Inequalities and multiple discrimination in access to and quality of healthcare. 2013 (https://op.europa.eu/en/publication-detail/-/publication/d50d3be5-c88b-4b46-af0e-debf4d1106ba/language-en).