Research

Ethical concerns around use of artificial intelligence in health care research from the perspective of patients with meningioma, caregivers and health care providers: a qualitative study

Melissa D. McCradden PhD, Ami Baba MRes, Ashirbani Saha PhD, Sidra Ahmad, Kanwar Boparai, Pantea Fadaiefard HBSc, Michael D. Cusimano MD PhD

Abstract

Background: As artificial intelligence (AI) approaches in research increase and AI becomes more integrated into medicine, there is a need to understand perspectives from members of the Canadian public and medical community. The aim of this project was to investigate current perspectives on ethical issues surrounding AI in health care.

Methods: In this qualitative study, adult patients with meningioma and their caregivers were recruited consecutively (August 2018–February 2019) from a neurosurgical clinic in Toronto. Health care providers caring for these patients were recruited through snowball sampling. Based on a nonsystematic literature search, we constructed 3 vignettes that sought participants' views on hypothetical issues surrounding potential AI applications in health care. The vignettes were presented to participants in interviews, which lasted 15–45 minutes. Responses were transcribed and coded for concepts, frequency of response types and larger concepts emerging from the interview.

Results: We interviewed 30 participants: 18 patients, 7 caregivers and 5 health care providers. For each question, a variable number of responses were recorded. The majority of participants endorsed nonconsented use of health data but advocated for disclosure and transparency. Few patients and caregivers felt that allocation of health resources should be done via computerized output, and a majority stated that it was inappropriate to delegate such decisions to a computer. Almost all participants felt that selling health data should be prohibited, and a minority stated that less privacy is acceptable for the goal of improving health. Certain caveats were identified, including the desire for deidentification of data and use within trusted institutions.

Interpretation: In this preliminary study, patients and caregivers reported a mixture of hopefulness and concern around the use of AI in health care research, whereas providers were generally more skeptical. These findings provide a point of departure for institutions adopting health AI solutions to consider the ethical implications of this work by understanding stakeholders' perspectives.

Artificial intelligence (AI) holds immense promise for health care.[1–3] The field is evolving rapidly owing to increased computing capacity, availability of data and the widespread adoption of electronic health records in hospitals. However, current trends in big data use have brought about ethical concerns regarding accountability, responsibility and trust, among others. The views of the public are essential to supporting institutions' approaches to adopting AI as well as guiding important education initiatives that may be crucial to maintaining the public's support and trust. However, the speed of progress and potential for benefit of the technology are are mired by ethical controversies surrounding the use of AI more broadly that may undermine public trust in this technology.[4]

Public perceptions regarding health data use for research are well characterized,[5–13] but limited work specific to public perceptions of AI has been noted.[14–16] Consistently across the globe, members of the public value the benefit to be gained from medical research but are concerned about the privacy of personal health data. Even before the era of big data, people described concerns about the use of health data for research, particularly as the data are made available to more individuals and groups. Paprica and colleagues[12] recently conducted a focus group on the use of health data with Canadian stakeholders and identified a strong suspicion of private industry, which had been noted by other investigators.[7,10,11] Similarly,

Kim and colleagues[13] discovered that it is important to patients whom their health information and biospecimens are shared with for research purposes; particular hesitance was observed in sharing data with for-profit institutions.

In the present study, we sought to expand the scope of inquiry provided by this prior work by using vignettes to elicit perspectives on a nonexhaustive set of ethical concepts that are central to AI applications in health care. We conducted qualitative interviews with patients, caregivers and health care providers to investigate their perspectives on ethical considerations of AI-enabled research.

Methods

Setting and design

Patients and caregivers were recruited consecutively (August 2018–February 2019) at St. Michael's Hospital, Toronto, in the senior author's (M.D.C.) neurosurgical clinic as part of a larger study focused on quality of life among patients with meningioma.[17] Patient eligibility criteria were diagnosis of meningioma, recent (2008–2018) neurosurgical or neuro-oncologic intervention and capacity to consent to research. Caregivers included spouses, adult children, relatives and friends who had accompanied the patient to at least 1 clinic appointment. Exclusion criteria were age less than 18 years and not fluent in English (given the complexity of the interview content). Health care providers caring for these patients were recruited through snowball sampling. No prior relationship existed between the participants and the interviewers, with the exception of 2 health care providers who had collaborative relationships with the primary investigator (M.D.C.). Participants were aware that the interviewers were part of a research group conducting AI work, which contributed toward the rationale of the present study.

After introductions by a member of the patient's circle of care, interviews were conducted in a private clinic room by a postdoctoral fellow (M.D.M.) or a research assistant (A.B.) (both female). The patient's caregiver was present in 2 cases. Interviews included collection of baseline demographic information and presentation of 3 vignettes (described below), and lasted about 15–45 minutes. All participants provided written informed consent. Interviews were audio-recorded and transcribed verbatim with consent; in cases in which consent was denied, the interviewer took detailed notes. No repeat interviews were conducted with any of the participants, and they did not see their interview transcripts.

Development of vignettes

To guide vignette development, we searched academic databases (PubMed, Medline, JSTOR and PsycINFO) and performed a Google search to identify a set of ethical principles prioritized consistently for health AI (Appendix 1, available at www.cmajopen.ca/content/8/1/E90/suppl/DC1). The final set included informed consent, privacy, confidentiality, responsibility, accountability, unintended consequences or harms, trust and public engagement (Table 1). We assessed perspectives around these ethical principles through 3 scenar-

ios: data-driven approaches to health care research, use of machine learning in clinics and commercialization of data (Appendix 1). Scenarios were trialled for comprehension and relevance to the intended ethical concept through 3 rounds of feedback with the research team and health care colleagues.

Participants were told that these scenarios were examples of realistic but hypothetical AI-enabled research. They were asked about their current knowledge of AI before the vignettes were described. After each scenario, participants were asked for their opinions and how they thought characters in the vignette would react or feel. Interviewers refrained from providing additional information beyond the details identified in the interview script (Appendix 1).

Data analysis

Directed content analysis (M.D.M., A.B., P.F.) allowed data interpretation under the umbrella of our predefined ethical concepts.[18] M.D.M. has formal education in empirical bioethics, including qualitative methodology, and has published qualitative work previously. A.B. has been trained in qualitative methodology. Open-ended questions were codified based on prevailing reasoning for the answer(s) given to a particular question. Closed-ended responses were categorized as yes (fully positive), no (fully negative), unsure (in-between positive and negative) or unknown, and justifications and reasons were noted.

Previous work indicated that thematic saturation is reached with 12 interviews.[19] The emerging themes were presented by the 2 interviewers for discussion with the primary investigator (M.D.C., who has done qualitative work in the past) and team members, who together decided when an acceptable level of saturation had been reached.

No coding software was used, and the data were managed with Microsoft Excel 2016.

Table 1: Ethical principles prioritized consistently for health artificial intelligence	
Concept	Definition
Consent	Agreement given free from coercion or undue influence having understood the benefits and risks
Privacy	Control over one's personal interests (e.g., personal health information)
Confidentiality	Obligation of institutions to safeguard entrusted information
Responsibility	Taking ownership of a decision
Accountability	Assigning blame, answerability, liability, proper accounting
Unintended consequences/ harms	Outcomes unforeseen, generated without purposeful action
Trust	Reliability, consistency in words and actions, guardianship
Public engagement	Supporting the meaningful participation of members of society

Ethics approval

Ethics approval was granted by the Unity Health Toronto Research Ethics Board.

Results

Of the 19 patients invited, 1 declined participation. All 7 caregivers and 5 health care providers invited agreed to participate. The participants' demographic characteristics are shown in Table 2. None had formal experience with AI systems or methodologies.

Providers' responses with highly consistent with each other. Patients and caregivers expressed divergent opinions on many issues and offered a range of different views. Overall, responses reflected a sense of uncertainty about what the "right" course of action should be in many circumstances (Table 3). Representative participant quotes concerning the ethical concepts are presented in Table 4.

Conditions of use of data for health care research

There was nearly unanimous agreement that health data are a valuable resource that can be directed for the purpose of improving health and disease treatment through research, but disagreement as to the threshold for requiring consent for their use. Many of those who advocated for consent initially felt that, in an urgent, disastrous situation (e.g., disease outbreak), the circumstances would be sufficiently compelling to warrant an "accelerated process" (participant 18–008, provider) or complete bypassing of consent. Many advocated for disclosure of health data use nonetheless, through social media, telephone calls, text messages or other media.

Most participants cited deidentification as a satisfactory condition for nonconsented use of health data for research. When asked about what deidentification meant, respondents agreed with removing any or all of name, social insurance number, date of birth, address or health care number, as prompted by the interviewer. These perspectives were connected explicitly with the use of data by researchers in health care for the purpose of improving medical care.

Deference to computer outputs

A minority of respondents readily accepted the idea that an output from a "computer" should allocate patients to treatment or no treatment based on a prediction from a computer regarding their probability of benefiting. The lone provider who agreed with this idea likened this to the obligation to not offer treatments that are unlikely to benefit a patient. Those who resisted this notion appealed to fairness or equality ("trying is more important" [participant 18–008, provider]), fair opportunity ("everyone deserves the chance to be treated" [participant 18–017, provider]), evidential uncertainty ("should do more research" [participant 18–015, caregiver]) and individual factors influencing prognosis. All but 1 provider rejected the notion of allocation of treatment by AI, appealing to the need for these decisions to be made collaboratively with patients.

Concerning the vignette that revealed there had been a mistake in the computer system, many participants declared that they had expected such a mistake, and nearly all were accepting of the notion that mistakes happen. Participants almost universally supported disclosure (although 1 patient disagreed, fearing repercussions to the algorithm developers) and reparations, including lawsuits ("they should suck it up and pay" [participant 18–007, caregiver]) and efforts to financially compensate and medically treat patients who

Table 2: Participant demographic characteristics

Characteristic	No. (%) of participants*		
	Patients n = 18	Caregivers n = 7	Health care providers n = 5
Gender			
Female	10 (56)	6 (86)	5 (100)
Male	8 (44)	1 (14)	0 (0)
Age category, yr			
20–30	0 (0)	1 (14)	1 (20)
31–40	2 (11)	0 (0)	1 (20)
41–50	3 (17)	2 (29)	2 (40)
51–60	5 (28)	1 (14)	0 (0)
61–70	4 (22)	2 (29)	1 (20)
71–80	1 (6)	0 (0)	0 (0)
81–90	3 (17)	0 (0)	0 (0)
Not disclosed	0 (0)	1 (14)	
Age, yr, mean	60.5	50.8	43.6
Highest level of education			
High school	2 (11)	1 (14)	0 (0)
College/university	13 (72)	4 (57)	2 (40)
Master's degree/doctorate	3 (17)	1 (14)	3 (60)
Not disclosed	0 (0)	1 (14)	
Ethnicity			
White	11 (61)	4 (57)	2 (40)
Black	1 (6)	0 (0)	1 (20)
Asian	3 (17)	1 (14)	1 (20)
Middle Eastern	1 (6)	0 (0)	0 (0)
Central American	0 (0)	1 (14)	0 (0)
European	2 (11)	1 (14)	1 (20)
Type of health care provider			
Neurosurgery resident	–	–	1 (20)
Medical administrative assistant	–	–	1 (20)
Nurse practitioner	–	–	1 (20)
Physiotherapist	–	–	1 (20)
International medical graduate	–	–	1 (20)

*Except where noted otherwise.

Table 3: Key participant perceptions regarding the use of artificial intelligence in health care research	
Concept	Participant perceptions
Consent	• Half of participants felt that each person must consent to allow his or her data to be used in research • Most were against private companies' obtaining data without individual consent • Most said that even deidentified data should not be sold to private companies without consent
Privacy	• Deidentification was believed by most participants to be the removal of name, social insurance number, date of birth, address, health card number • Some participants felt that the loss of privacy is an acceptable sacrifice for the prospect of benefit to the larger population
Confidentiality	• Most participants felt that conditions under which individuals provide consent ought to be respected (e.g., use of data for health research v. marketing purposes) • All providers felt that receiving health information is a privilege given its highly sensitive nature
Responsibility	• Some participants accepted allocation of health resources via computerized output • Half felt it inappropriate to delegate responsibility to computers • One provider likened delegation of responsibility to computers to inappropriate treatment; the other providers advocated for shared decisions
Accountability	• Some participants indicated media as a key mechanism for accountability • Some participants indicated skepticism that institutions and companies could be held accountable
Unintended consequences/harms	• Most participants accepted that mistakes happen • All stressed the need for transparency, disclosure and reparations • Some felt that transparency and publication prevent others from repeating the mistake
Trust	• Most participants felt that health care institutions are highly trusted organizations • Most participants felt that physicians and other health care providers are entrusted with carrying out research with health data
Public engagement	• Some participants felt that the public had a duty to be involved in research in some way • Most were unsure how specifically to have a voice in medical research

Table 4: Illustrative participant quotes regarding ethical issues	
Issue	Representative quote
Protection of health data	The world that we live in, there's all kinds of access to information even though it's protected, but you hear all kinds of scenarios where sensitive information gets leaked. So yeah, I would have some concerns. (Participant 18–042, patient)
Skepticism regarding accountability mechanisms	As a member of the public, my opinion doesn't count. (Participant 18–004, patient)
Allocation of treatment by computers	It's ethically incorrect, as you are picking and choosing who gets treatment. You need to give them options and have conversations with the patients. (Participant 18–008, provider)
Allowing sale of data to private industry	You should have to give up some [privacy]. … You want to be cured and [the company is] providing you with this cure, so you balance it out. (Participant 18–012, caregiver)
Computer-based predictions	Before [the brain tumour], I might [have said] yes, because I would say … it's the survival of the fittest. … But you can never underestimate the fight … in a person, even with a disease. And [a patient] can far surpass the expectations that are set out in these kinds of statistics. (Participant 18–001, patient)
Trust and confidentiality	I think … in a democratic society, for members of the public to have faith in the health care system … individuals need to believe that what they believe to be confidential is held confidential, and not shared. But also for me to have confidence in health care systems, I have to believe that leaders in health care systems will make decisions for the greater good of people, right? (Participant 18–032, patient)
Health data v. other data	It's a privilege to be told this information — patients don't even tell their family what they tell us (Participant 18–054, provider)

were excluded from treatment. Some were less forgiving ("fire them and hire new researchers" [participant 18–049, patient]). When asked who was responsible for the mistake, most participants pointed to those who developed the algorithm, with a few specifically blaming the people who input the data into the computer. One participant said that the person most in charge was responsible for the outcome. One provider described the need to publish and report the negative results so that others would not repeat the mistake.

Secondary use or sale of data

Most participants felt strongly that selling health data to private companies should be prohibited entirely. The few who disagreed argued that loss of privacy is an acceptable sacrifice for the prospect of benefit to the larger population, indicating that, as long as the product being developed would help people and adverse effects were minimal, selling health data was justified. Others described the difficulty in not knowing what kind of product would be developed; 1 participant noted "every company thinks [it's] honourable, but it depends on your perspective" [participant 18–002, patient]). Overwhelmingly and regardless of their view, participants advocated for transparency about how health data would be used, communicated openly by a trusted institution or custodian of health information.

No health care providers felt that selling either identifiable or deidentified data was appropriate. They perceived that selling data conflicted with the responsibilities of health data custodians. One provider described patients as a "vulnerable population," as patients are eager to support any endeavour purported to help others with the same disease, even if they know they themselves will not benefit directly (participant 18–010, provider). The idea that research might be able to "find a cure" was echoed repeatedly in this context by patients and caregivers, seemingly supporting providers' views.

Trust and public engagement in research

Patients and caregivers reported a high level of trust in health care institutions with regard to ethical practices and acting responsibly vis-à-vis health data by following regulations designed to protect the public. When asked about a duty to participate in research specifically through allowing use of their health data, a few participants stated that people had a duty to allow such use for the specific purpose of researching health-related problems, whereas others indicated no one had such a duty. Nearly all participants who did not express a yes or no answer indicated that they personally felt a sense of duty to contribute their data to research but that not everyone would agree, and individuals' wishes should be respected. Others described a duty only if the research involved deidentified data and no potential harms to participants.

Several participants described a morally significant difference between data obtained from social media versus health data. All providers stated that health data were special, whereas most patients and caregivers indicated that, in modern society, people are now aware of the consequences of smartphone use, resulting in the minimization of privacy concerns. Despite a perception that data sharing is now inevitable, most participants clearly indicated discomfort with the lack of transparency regarding how their data were being used.

Interpretation

Our participants generally did not have substantively different perceptions regarding the use of AI in health research compared to previously explored notions of health data use and research.[5–13] Exploring machines as "decision-makers," however, elicited a range of opinions, with some participants being at ease with allowing such decisions to be delegated to machines and a majority expressing skepticism.

Our participants endorsed the notion that the broad use of health data as a resource to improve health[11,12,20] poses risks to personal privacy, in keeping with a previous report.[5] Patients derived a sense of altruism in providing their data, which contrasted with the feeling of powerlessness in having a brain tumour; this notion corresponded with 1 provider's statements that patients may constitute a vulnerable group. Vulnerable groups, however, are not uniform;[21–23] although our patient participants strongly supported research, vulnerability from a racialization or socioeconomic perspective often connotes distrust.[9] The willingness to engage in health data research that drives AI is likely modified by the disproportionate risks inherent to AI that are carried by various marginalized populations.[24–26]

McDougall[27] speculated that AI may disrupt patient autonomy. We found limited endorsement of deference to computerized outputs in our sample. Xu and colleagues noted that some people would blindly trust a robot to guide them through a rehabilitation protocol.[14] However, those authors used an interactive robot that approximated a human interaction, whereas we described the guidance coming from a "computer" (i.e., nonhumanoid). Most (4/5) of our provider participants indicated that treatment decisions require conversations with patients and families. Even among tech-savvy youth seeking treatment for highly stigmatized conditions in a qualitative study, there remained a strong preference for interacting with health care providers to discuss health issues.[16]

A particular challenge for health AI involves the often-needed collaboration with private industry to develop solutions. Like Paprica and colleagues,[12] we found a generally more negative or mixed reaction to sharing health data with private companies. Our participants sharply contrasted the use of data to improve health with prioritizing profit-making.[7,10–12] Patients' trust in health care institutions compels providers to retain a strong understanding of the social licence[12] surrounding health data use as AI is integrated into health care.

Future studies may extend these findings by soliciting views from AI-knowledgeable people. In addition, although our sample was not homogenous, it was selective in that it included people with high levels of health care interactions. It will no doubt be important to capture the views of a more diverse group of people with varying levels of health care interactions.

Limitations

Our study's findings may have limited generalizability given the population (patients with meningioma and their caregivers and health care providers) and participants' extensive involvement with the health care industry. The health care providers interviewed were selected by convenience sampling and so may not be representative of clinicians generally. However, the responses they provided were consistent with prior work, which suggests that many central concepts such as appropriate use of health data and consent may be inherent to clinicians' professional duties and less likely due to sampling bias.

It is possible that more detailed views may have been missed because we left the scenarios somewhat vague. This was consistent with the study aim, which was to provide an initial glimpse into views on health AI. We also intended to avoid getting overly complex about details that are not yet formalized with regard to AI's broader adoption in health care. To be highly specific at this juncture would be premature.

Conclusion

We highlight initial perspectives surrounding the use of AI in health care research among AI-naive patients, caregivers and health care providers at a large urban hospital. We found a mixture of hope and skepticism regarding the use of AI. The findings reflect previous work citing tensions between privacy and potential benefit. Although there was broad support for the use of AI in health research, this study identified certain caveats, including the desire for deidentification of data and use within trusted institutions, with the goal of contributing toward the improvement of health.

References

1. Hinton G. Deep learning — a technology with the potential to transform health care. *JAMA* 2018;320:1101-2.
2. Israni ST, Verghese A. Humanizing artificial intelligence. *JAMA* 2019;321:29-30.
3. Naylor CD. On the prospects for a (deep) learning health care system. *JAMA* 2018;320:1099-100.
4. Gibney E. The battle for ethical AI at the world's biggest machine-learning conference. *Nature* 2020;577:609.
5. Robling MR, Hood K, Houston H, et al. Public attitudes towards the use of primary care patient record data in medical research without consent: a qualitative study. *J Med Ethics* 2004;30:104-9.
6. Haddow G, Bruce A, Sathanandam S, et al. "Nothing is really safe": a focus group study on the processes of anonymizing and sharing of health data for research purposes. *J Eval Clin Pract* 2011;17:1140-6.
7. Lehnbom EC, Brien JE, McLachlan AJ. Knowledge and attitudes regarding the personally controlled electronic health record: an Australian national survey. *Intern Med J* 2014;44:406-9.
8. Grande D, Mitra N, Shah A, et al. The importance of purpose: moving beyond consent in the societal use of personal health information. *Ann Intern Med* 2014;161:855-62.
9. Bansal G, Zahedi F, Gefen D. The impact of personal dispositions on privacy and trust in disclosing health information online. *Decis Support Syst* 2010;49:138-50.
10. Willison DJ, Swinton M, Schwartz L, et al. Alternatives to project-specific consent for access to personal information for health research: insights from a public dialogue. *BMC Med Ethics* 2008;9:18.
11. Gaylin DS, Moiduddin A, Mohamoud S, et al. Public attitudes about health information technology, and its relationship to health care quality, costs, and privacy. *Health Serv Res* 2011;46:920-38.
12. Paprica PA, de Melo MN, Schull MJ. Social licence and the general public's attitudes toward research based on linked administrative health data: a qualitative study. *CMAJ Open* 2019;7:E40-6.
13. Kim J, Kim H, Bell E, et al. Patient perspectives about decisions to share medical data and biospecimens for research. *JAMA Netw Open* 2019;2:e199550.
14. Xu J, Bryant DG, Howard A. Would you trust a robot therapist? Validating the equivalency of trust in human-robot healthcare scenarios. *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2018 Aug. 27–31; Jiangsu Conference Center, Nanjing, China:442-7. Available: https://ieeexplore.ieee.org/abstract/document/8525782/ (accessed 2019 Aug. 19).
15. Hengstler M, Enkel E, Duelli S. Applied artificial intelligence and trust — the case of autonomous vehicles and medical assistance devices. *Technol Forecast Soc Change* 2016;105:105-20.
16. Aicken CRH, Fuller SS, Sutcliffe LJ, et al. Young people's perceptions of smartphone-enabled self-testing and online care for sexually transmitted infections: qualitative interview study. *BMC Public Health* 2016;16:974.
17. Baba A, McCradden MD, Rabski J, et al. Determining the unmet needs of patients with intracranial meningioma — a qualitative assessment. *Neurooncol Pract* 2019 Oct. 29. doi: 10.1093/nop/npz054.
18. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;15:1277-88.
19. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 2006;18:59-82.
20. Hastings TM. Family perspectives on integrated child health information systems. *J Public Health Manag Pract* 2004;10(Suppl):S24-9.
21. Levine C, Faden R, Grady C, et al.; Consortium to Examine Clinical Research Ethics. The limitations of "vulnerability" as a protection for human research participants. *Am J Bioeth* 2004;4:44-9.
22. Macklin R. Bioethics, vulnerability, and protection. *Bioethics* 2003;17:472-86.
23. Rogers W, MacKenzie C, Dodds S. Why bioethics needs a concept of vulnerability. *Int J Fem Approaches Bioeth* 2012;5:11-38.
24. Char DS, Shah NH, Magnus D. Implementing machine learning in health care — addressing ethical challenges. *N Engl J Med* 2018;378:981-3.
25. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366:447-53.
26. Schulz A, Caldwell C, Foster S. "What are they going to do with the information?" Latino/Latina and African American perspectives on the Human Genome Project. *Health Educ Behav* 2003;30:151-69.
27. McDougall RJ. Computer knows best? The need for value-flexibility in medical AI. *J Med Ethics* 2019;45:156-60.

Affiliations: Division of Neurosurgery (McCradden, Baba, Saha, Boparai, Fadaiefard, Cusimano), St. Michael's Hospital, Unity Health Toronto; Dalla Lana School of Public Health (Cusimano), University of Toronto, Toronto, Ont.