July 2018

ETHICS AND AI FOR GOOD HEALTH

SYMPOSIUM REPORT





Symposium Overview

Artificial intelligence (AI) has been described as the "fourth industrial revolution" with far-reaching implications globally, including in health, and health care. The *Ethics and AI for Good Health Symposium*, hosted by the Joint Centre for Bioethics (JCB) on June 11, 2018, is the inaugural workshop of the "Collaboration on Artificial Intelligence for the Public's Health", a new partnership of the Dalla Lana School of Public Health, Public Health Ontario, and the Vector Institute for artificial intelligence.

In this Symposium Report, we summarize the symposium proceedings, report key findings about ethics and AI for health based on the expert contributions of Symposium panelists and participants, and identify immediate next steps. The development of this report took into account the current projects and perspectives of the field experts who attended and presented at the *Ethics and AI* for Good Health Symposium.

SYMPOSIUM OBJECTIVES

The objectives of the symposium were to: (1) explore ethical, social, and legal issues related to AI for health; (2) facilitate cross-disciplinary exchange about the science and application of AI for health; and, (3) foster collaboration for research and education within our local academic health ecosystem.

PARTICIPANTS

The symposium was attended by 70 participants including individuals affiliated with the Global Network of WHO Collaborating Centres for Bioethics as well as other local and international experts representing health professionals, data scientists, legal scholars, ethicists, epidemiologists, global health, and population health specialists amongst others. We would like to extend our thanks for their contribution and insight which made this review possible.

FORMAT AND METHOD

This one-day symposium involved a series of panel presentations and interactive group discussions on ethics and AI for health in three domains: health care; population health, health systems, and public health practice; and global health (see Appendix 1: Symposium Agenda).

This symposium report was prepared by the Ethics & AI for Good Health Project Team at the University of Toronto's Joint Centre for Bioethics (JCB). Special thanks to Neha Malhotra and Jia Ce Cai for drafting the report. Funding support was provided by a CIHR planning grant. For more information about the JCB's work on ethics & AI, contact Dr. Jennifer Gibson at jcb.director@utoronto.ca.

A key input to the symposium was a Backgrounder document prepared by the JCB team, which summarized six ethics domains relevant to AI for health based on a grey literature review (see Appendix 2: Symposium Backgrounder). Each panel included three experts who were either using or planning to use AI. These experts provided insight on the opportunities and challenges associated with applying AI in these health domains, including any potential ethical issues. Following each panel, symposium participants were invited to discuss the presentations with others in small groups and record any questions or comments for the panel on post-it notes provided. Each table was then asked to pose one question to the panel. All post-it notes were collected and incorporated in the analysis of the symposium findings.

Symposium Proceedings

OPENING REMARKS

Adalsteinn Brown, Dean, Dalla Lana School of Public Health Alison Paprica, Vice President, Health Strategy and Partnerships, Vector Institute for artificial intelligence

This symposium aimed to bring together expertise from different domains at the intersection of health, ethics, and AI. We intended for the day to be a "meeting of minds" where participants came with a mixture of questions and concrete ideas on how to navigate this domain. Our hope is that each of the participants left the symposium with new insights and collaborative opportunities.

Steini Brown, Dean of the UofT Dalla Lana School of Public Health, noted that AI is transformative for health, health care, and public health. As such, we must reflect on the ethical implications of AI to ensure we do not reinforce systemic biases that further exacerbate existing health inequities. It is important to have a strong ethical framework as AI is integrated into our health systems.

Alison Paprica, VP, Health Strategy and Partnerships at the Vector Institute for artificial intelligence, noted in her opening remarks that while computers can generate insights more than humans can through traditional methods, how we employ AI requires further consideration. Ms. Paprica urged the audience to reflect on how academia can involve the public in data intensive research to establish a social license for AI. As a platform to disseminate recommendations, the Vector Institute stands ready to implement ideas that come out of the symposium discussions.

OPENING ADDRESS: ETHICS AND AI FOR HEALTH: WHAT DO WE KNOW? AND WHY DOES IT MATTER?

Ross Upshur, Professor and Head, Division of Clinical Public Health, DLSPH Jennifer Gibson, Director, Joint Centre for Bioethics Moderator: Erica Di Ruggiero, Director, Office of Global Public Health Education and Training, DLSPH

Ross Upshur set the stage for the symposium in an opening address on what we know thus far on the use of AI for health and care. Recognizing that ethics and technology have co-evolved, Upshur noted that we can learn "best practices" from previous ethical frameworks. While AI seems to be a relatively new technology, we must reflect on its supposed novelty. Are there issues that have been raised by technologies such as drugs, transplants, and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) that are analogous to AI? We must consider what makes AI different from its technological predecessors. AI will no doubt require some adaptation and innovation in ethics and governance approaches, and such guidance should promote the principles of inclusivity.

Following this discussion, Jennifer Gibson presented some preliminary findings from a grey literature search, which prompted some feedback and discussion from the audience. Six themes were highlighted in the Backgrounder (see Appendix 2): (1) data quality and use; (2) equity and the digital divide; (3) governance; (4) collaboration; (5) redefining health care practice; (6) the future of work (see figure below).



In response to these thematic areas, some participants suggested redefining "collaboration" as it may not fully capture the tensions between the different roles and contributions that stakeholders may have as a result of differing values. Furthermore, the "ethical design of AI" was cited as an area that requires further exploration under its own thematic area to understand how we can train AI to make ethical decisions. Lastly, communication of data sets and AI to different communities, patient populations, health care professionals, and tech experts needs to be addressed to ensure that we have a common vocabulary for public discourse. This presents an opportunity for intercultural dialogue within an increasingly globalized context.

PANEL 1: AI IN HEALTH CARE

Frank Rudzicz, Scientist, University Health Network and Vector Institute Sally Bean, Director, Ethics Centre, Sunnybrook Health Sciences Centre Ken Goodman, Director, University of Miami Ethics Programs Moderator: Jennifer Gibson

Frank Rudzicz is a computer scientist using speech recognition and artificial intelligence in applications designed for individuals with special speech characteristics. Reflecting on his work with AI in health care, Rudzicz highlights that "[the risk of AI] is not disobedient AI, but instead, obedient AI mixed with lazily defined objectives". In order to ensure the responsible use of AI in health, we must be cognizant of potential biases represented in data that can be translated into biased algorithmic decisions.

Sally Bean is a practicing healthcare ethicist with a professional background in law, who is interested in responsible innovation and ethical and societal values incorporated into AI. Citing the need for ethicists to be key allies in AI innovation, Bean calls for us to develop trust amongst end users of technology to avoid the current situation with vaccines. To prevent AI hesitancy, as a result of complacency, inconvenience, and lack of confidence in technology, Bean suggests incorporating principles of fairness, accountability, and transparency/trust into responsible innovation practices.

Ken Goodman is a philosopher and expert in the ethics of health informatics. He describes new technologies, like AI, as rich sources of ethical concern, tension and progress. Goodman highlights four overarching concerns for AI: altering traditional relationships, technological abuse, depersonalization, and unintended consequences. While there is a possibility of improvement in health outcomes, it is imperative we consider the possible erosion or elimination of traditional relationships in the health care setting.

PANEL 2: AI IN POPULATION HEALTH, HEALTH SYSTEMS, AND PUBLIC HEALTH PRACTICE

Natasha Crowcroft, Chief, Applied Immunization Research and Evaluation, Public Health Ontario (PHO) Muhammad Mamdani, Director, Li Ka Shing Centre for Healthcare Analytics Research and Training (LKS-CHART) Laura Rosella, Assistant Professor, DLSPH Moderator: Ross Upshur

Natasha Crowcroft is a public health physician and scientist whose primary areas of research are in infectious disease and immunization. She outlined how machine learning technologies might, in principle, be leveraged to complex public health problems particularly related to infectious diseases. However, she noted that this would require a collaborative interdisciplinary model among public health researchers and AI scientists. Her work on simplifying complex public health data through language processing and machine learning surfaced issues such as under-trained staff in AI technology. She suggested that a collaboration between researchers and AI experts will make data processing more efficient while also decreasing data biases and inequities. As Crowcroft stated: "we need to spend time together to learn each other's languages".

Muhammad Mamdani, Director of the LKS-CHART, focussed on AI developments in hospitals. His project on clinical data systems at St. Michael's Hospital found the unintended consequences of using algorithms in patient treatment plans. For example, AI could guide clinical decisions that may extend a patient's life, but with extremely poor quality. Mamdani concluded that AI still has gaps which need to be filled in order to produce beneficial results for all.

Laura Rosella, an epidemiologist who founded the Population Health Analytics Lab at the DLSPH, discussed the predictive benefits of AI for public health. Rosella's research showed that AI and machine learning offer opportunities to integrate complex and unstructured data sources on the broader social and environmental determinants of health into population risk assessment. As well, since data that is more representative of communities yields less measurement errors, she emphasized that we must overcome differences in data access and AI knowledge so that everyone can benefit from AI.

PANEL 3: AI FOR HEALTH IN A GLOBALIZED WORLD

Andreas Reis, Technical Officer, Global Health Ethics, World Health Organization Erica Di Ruggiero, Director, Office of Global Public Health Education and Training, DLSPH

Amy Fairchild, Professor and Associate Dean, Texas A&M School of Public Health Moderator: Solly Benatar

Andreas Reis from the WHO who has worked extensively in global public health policy and ethics, presented on various projects and programs relating to big data and AI. With new leadership, the WHO is currently implementing a five year program emphasizing new technologies such as big data and AI while putting ethics at the heart of its mandate. We learned that the WHO is concerned with digital inequalities, and thus seeks expertise to guide and support low- and middle-income countries to benefit from big data and AI. Additionally, there needs to be ethical frameworks both nationally and internationally to ensure safe and fair use of AI. Through these programs, governments' health related goals could be realized with the help of AI technology.

Erica Di Ruggiero, whose research examines how evidence affects global policy agendas related to employment and health equity, focussed her presentation on the changing nature of work in a globalized, AI-centric world. While new employment opportunities are likely to emerge with the use of AI, these may displace women, who typically deliver care, and increase demand for well-paid, male-dominated STEM positions. There is also an increase in less well-paid jobs and "gig" work to be expected. For the future of work, what is the obligation of companies and governments to workers made redundant by AI?

Amy Fairchild, a public health historian, proposed that there might be a new hybrid between private and public institutions to better harness the benefits of AI technology. Specializing in surveillance and privacy issues, Fairchild stated that the social and health inequities that may emerge within and between countries existing in AI technology are threatening its capacity to expand. Her presentation showed that there is a moral obligation for high income countries to support countries that lack infrastructure to collect data.

CLOSING PLENARY: ETHICS AND AI FOR HEALTH: WHERE DO WE GO FROM HERE?

The closing plenary discussion opened the floor for participants to reflect on the day's proceedings. Symposium participants underscored the need for broader public consultation on AI and its uses in relation to health. They also emphasized the need for a better understanding of patient and public expectations with respect to privacy in this context. Another emerging theme was trust, which participants tied to ensuring transparency in the development and application of AI and fostering a common language across disciplines and actors involved in AI and health. Lastly, there is also a need for robust models of collaboration emphasizing partnership with strong ethical governance and accountability between and within public and private sector institutions.

Key Findings

CROSS-CUTTING THEMES

Accountability

The application of AI challenges current understandings of accountability. Within health care, who ought to be responsible should adverse events results from the usage of AI technology? For the public health sector, employer-employee relationships in digital platform-mediated gig work differ from traditional modalities of work. As a result, there is a lack of accountability for worker protection. Lastly, we need ethical frameworks both nationally and internationally to support responsible innovation and accountability. We must ensure that both public and private sector institutions remain accountable to using AI for good and responding to societal needs and not further widen inequities.

Trust

Numerous participants and panelists highlighted maintaining trust in AI as a foreseeable challenge. WHO international consultations on ethics, big data and AI report that the role of trust and governance is one that continues to be identified by experts across the globe. There are many facets of trust in response to the usage of AI in health. There must be trust in the quality of algorithms, trust in the data used in the algorithm, and trust in clinicians using AI outputs. In order to maintain trust in the development and application of AI technology for health, there must be a willingness to engage with the public. Considering AI is a broad term often misrepresented in popular media, there is a need for greater public understanding of what AI is and its usage in health.

Privacy

The issue of privacy from a clinical and data collection perspective was heavily discussed throughout the symposium. As privacy strongly relates to trust, both presenters and attendees alike had strong opinions to prioritize data confidentiality so that patients are neither discouraged to participate nor fearful that their personal data is accessed by third party sources. Additionally, privacy laws to protect not only patients but also the general public were suggested to adhere to ethical research standards if extensive data collection or data mining is performed on both a national and international level. Considering its connection public trust, AI development should continue to emphasize privacy.

PANEL-SPECIFIC THEMES

1. Health Care

Bias in data

Artificial intelligence is dependent on big data to learn and train algorithms. As such, it can easily inherit any biases that may exist as a result of sampling errors. While demographic data like skin colour may be clearly identified, there may be biases that are unknown to us that we can capture through AI. Given the potential for bias in data, there are differing and conflicting views on how these biases can be overcome. We need to understand if there is a threshold at which we have a sufficient amount of data to minimize concern of biases being represented in AI.

Responsible use of AI

The responsible use of AI can be understood in terms of how we ought to use AI. In order to understand responsible innovation and application of AI, we must ask who decides what constitutes the reasonable and appropriate use of AI, and whether current frameworks and processes are up to the task. Furthermore, various regulations and governance frameworks result in differing ethical implications of AI depending on whether you represent public or private sector institutions. It is important to universally govern AI to ensure it is being used appropriately.

Changes in health care relationships

Alterations in the trust-based human relationships in health care are a growing concern in the face of AI. The changing role of humans in health care is also an issue. Another challenge identified is the shifting responsibility of care and monitoring from health professionals to patients. We must consider whether this new dynamic in health care practice will disadvantage patients who do not have the time, resources or access to technology.

2. Population Health, Health Systems, and Public Health Practice

Shared understanding through collaboration

Fostering collaboration between professionals of different fields will be important to the progress of AI technology as each field brings unique perspectives to the table. However, collaboration can be hindered by differences in knowledge regarding AI as well as variations in resources. Thus, we need a concrete and specific way to engaged with AI technology so that discrepancies in approaches can be mitigated while ensuring centres without adequate resources and capacity are receiving the support they need.

Algorithm development

Before AI can be utilized, there needs to be a developmental process for the algorithm upon which it is based. This process is crucial as the algorithm directly relates to AI output utilized for decision making in many different fields of health. The question currently lies in how trustworthy the algorithm is and how to best minimize bias and promote ethically conscious algorithms. By answering these questions, we will be able to structure algorithms to yield more reliable results.

Risk prediction through AI outputs

In population and public health, one of the major benefits of utilizing AI technology will be the ability to forecast future health risks on a global scale. An operation of this scale will require a substantial amount of data collected on not only health information but also environmental and socioeconomic factors. The large amount of data poses an issue as measurement errors can prevail. The question is how we can first obtain such diverse data while maintaining the integrity of the results.

3. Health in a Globalized World

Digital divide

As AI only reflects data with which it has been trained, it is important to input data which considers all communities from both high, middle- and low income countries in order to minimize bias in AI output. Doing so will mean overcoming the "digital divide" where certain countries do not have the technical framework to collect their own data. This is a prominent issue as under-represented groups will not be benefiting from AI. Therefore, during the advancement of AI technology, we also need to be cognizant of what infrastructure and supports will be required to support less resource-rich nations.

Work

With the widespread implementation of AI, there will be a drastic shift in the nature of work as many occupations will become automatized. As this process occurs, who will lead the ethical transition from the jobs which AI displaces towards new jobs which AI creates? We should consider how the displacement of jobs will affect the socio-economic status of citizens.

Public vs. Private

Since public and private sectors both have distinct advantages to offer, their combined forces will generate a greater capacity to develop AI technology at a faster and more effective rate. However, with the creation of a public and private hybrid for AI advancement, there are questions around who will be leading them and what will future regulations look like.

Next Steps

This symposium report provides an overview of the day's deliberations and a detailed analysis of the themes that emerged from the panel contributions and participant discussions. As part of the JCB's Ethics and AI for Health project, a scoping review of relevant peer and grey literature at the intersection of AI, ethics, and health is ongoing. The project findings are to be released by September 2018, which will inform two subsequent workshops of the "Collaboration on Artificial Intelligence for the Public's Health" in Fall 2018.



ETHICS AND AI FOR GOOD HEALTH

SYMPOSIUM AGENDA















8:00 AM	REGISTRATION AND BREAKFAST
	Location: 41 Willcocks St, Faculty Club, Dining Room
8:30 AM	WELCOME AND INTRODUCTIONS
	Jennifer Gibson, Director, Joint Centre for Bioethics
8:45 AM	OPENING REMARKS
	Adalsteinn Brown, Interim Dean, Dalla Lana School of Public Health
	Alison Paprica, Vice President, Health Strategy and Partnerships, Vector Institute for artificial intelligence
9:00 AM	TODAY'S GOALS AND OBJECTIVES
9:05 AM	ETHICS AND AI FOR HEALTH: WHAT DO WE KNOW? AND WHY DOES IT MATTER?
	Ross Upshur, Professor and Head, Division of Clinical Public Health, DLSPH
	Erica Di Ruggiero, Director, Office of Global Public Health Education and Training, DLSPH
	Jennifer Gibson, Director, Joint Centre for Bioethics
10:15 AM	BREAK

10:45 AM	PANEL 1: AI IN HEALTH CARE
	Frank Rudzicz, Scientist, University Health Network and Vector Institute
	Sally Bean, Director, Ethics Centre, Sunnybrook Health Sciences Centre
	Ken Goodman, Director, University of Miami Ethics Programs
	Moderator: Jennifer Gibson
12:00 PM	LUNCH
1:00 PM	PANEL 2: AI IN POPULATION HEALTH, HEALTH SYSTEMS, AND PUBLIC HEALTH PRACTICE
	Natasha Crowcroft, Chief, Applied Immunization Research and Evaluation, Public Health Ontario
	Muhammad Mamdani, Director, Li Ka Shing Centre for Healthcare Analytics Research and Training
	Laura Rosella, Assistant Professor, DLSPH
	Laura Rosella, Assistant Professor, DLSPH Moderator: Ross Upshur

2:30 PM PANEL 3: AI FOR HEALTH IN A GLOBALIZED WORLD

Andreas Reis, Technical Officer, Global Health Ethics, World Health Organization

Erica Di Ruggiero, Director, Office of Global Public Health Education and Training, DLSPH

Amy Fairchild, Professor and Associate Dean, Texas A&M School of Public Health

Moderator: Solly Benatar

3:45 PM ETHICS AND AI FOR HEALTH: WHERE TO FROM HERE?

Jennifer Gibson, Erica Di Ruggiero, Ross Upshur

4:15 PM NEXT STEPS AND CLOSING REMARKS



ETHICS AND AI FOR GOOD HEALTH

SYMPOSIUM BACKGROUNDER





Acknowledgements

AUTHORS

Neha Malhotra

Research Assistant, Joint Centre for Bioethics MPhil Candidate, University of Cambridge BHSc (Honours), Global Health Specialization, McMaster University

Kathleen Murphy

Research Assistant, Joint Centre for Bioethics BSc, St. Francis Xavier University MPH, University of Toronto

Special thanks to Dr. Jennifer Gibson, Dr. Erica Di Ruggiero, and Dr. Ross Upshur for their guidance and invaluable feedback.

Symposium Goals

The Ethics and AI for Good Health Symposium, hosted by the Joint Centre for Bioethics, is the inaugural workshop of the "Collaboration on Artificial Intelligence for the Public's Health", a new partnership of the Dalla Lana School of Public Health, Public Health Ontario, and the Vector Institute for artificial intelligence.

The inaugural symposium has three aims:

To explore ethical, social, and legal issues related to AI for health

2

1

To facilitate cross-disciplinary exchange about the science and application of AI for health

3

To foster collaboration for research and education within our local academic health ecosystem



Executive Summary

The Joint Centre for Bioethics launched an Ethics and AI for Health initiative in Fall 2017. Funded in part by a CIHR Planning Grant, the initiative aims to scope and synthesize the terrain of ethical issues associated with the development and application of AI in the health field. In this Backgrounder, we provide initial findings of the grey literature component of a full scoping review based on the research question: *What ethical issues have been identified in relation to AI in the field of health locally and globally?* This backgrounder summarizes broad themes emerging from the grey literature and provides an initial foundation for discussion of ethics of AI in health locally and globally.

DATA QUALITY AND USE

- How should potential bias in datasets and algorithms be addressed?
- How ought individual rights be balanced with the potential for broader health benefits of AI?
- What are the ethical implications of commodifying data for commercial use?

EQUITY AND THE DIGITAL DIVIDE

- What impact will AI technologies have on health and social inequities?
- How do we include and build capacity in countries that have yet to undergo mass digitization and lack sufficient resources to benefit from Al's potential?
- How do we ensure that AI innovation is responsive to the unique needs of countries across the globe?

GOVERNANCE

- How do we maintain public trust in the development and application of AI technology for health?
- What key ethical considerations ought to be incorporated into regulatory frameworks locally and globally?
- How might ethical governance frameworks support responsible innovation and accountability?



Executive Summary

COLLABORATION

- Whose perspectives ought to inform the agenda for AI research and development?
- What are promising models for collaboration between AI developers, ethicists, and social scientists to ensure technologies are always developed with the necessary foresight of its impact on humanity?
- How do we ensure that designer intent is consistently informed by the interests and values of those who will be most affected by these technologies?

REDEFINING HEALTH CARE PRACTICE

- What are the ethical and social implications of using AI technologies in health service delivery and clinical decision-making?
- Who is accountable if an adverse event results from using AI technology in a patient's care?
- If digital health apps will benefit patient care, how will digital and health literacy gaps be addressed to avoid creating new health inequities?

FUTURE OF WORK

- Are women and marginalized populations at greater risk of job displacement in the face of AI adoption in the health care sector?
- What is the responsibility of higher education institutions to better prepare future health professionals and researchers to work alongside AI technology?
- How do we integrate social sciences and humanities into the training curriculum for future developers and programmers?



Methodology

The Joint Centre for Bioethics (JCB) launched its Ethics and AI for Health project in Fall 2017 with funding support of a CIHR Planning grant and the Jus Innovation Fund, an endowed fund of the JCB. The goals of the project are to: i) To describe *key ethical issues and considerations* related to AI and health locally and globally, ii) To *foster interdisciplinary engagement* on ethics and AI for health, iii) To strengthen *local and international collaboration* in support of ethical AI in health policy and practice; and iv) to develop an *agenda* for interdisciplinary health research on ethics and AI.

A Scoping Review of the academic and grey literature is underway as part of this project. This Backgrounder reports preliminary findings of the grey literature scan. Grey literature¹ was retrieved from grey literature databases, including OAlster, Google Scholar, the Canadian Electronic Library, and the Canadian Institute for Health Information; a customized Google search to identify documents from think tanks, Canadian government, and non-governmental organizations; and targeted website searches of known organizations and institutions . A total of 800 records ² were retrieved and a random sample of 50% (n=400) were screened for inclusion. A record was included if it was written in English and if it addressed AI, health and ethical or social considerations. Of the 400 screened articles, 35 met the inclusion criteria and underwent full text review by two members of the research team. A thematic analysis was performed to identify key domains of interest at the intersection of ethics, AI, and health, which are reported here.

This Symposium Backgrounder reports preliminary findings of the grey literature scan as an input to Symposium discussions and future workshops of the 'Collaboration on AI for the Public's Health', a new partnership of the Dalla Lana School of Public Health, Public Health Ontario, and the Vector Institute for artificial intelligence. It is a green document that will evolve through these expert discussions locally and globally and the final Scoping Review analysis later this summer.

¹ Grey literature is defined as any literature that has not been published through traditional means, and is excluded from mainstream sources (Gerstein Science Information Center, 2018). For the purposes of this paper, we have excluded conference proceedings and conference papers given their overlap with the academic literature, however they will be included in the larger scoping review.

² These searches were for records published between 2015-2018. The grey literature databases and Google searches were conducted between May 10- May 13 2018, and the targeted website searches occurred between April 25-May 19, 2018.

Data Quality & Use

The responsible and ethical application of AI technology is closely intertwined with good data practices such as adequate representation of socioeconomically and ethnically diverse populations and maintaining the rights and liberties of the people being served.

First, the success of AI technology is largely dependent on the quantity and quality of its data feed. In order to ensure that AI decisions are reliable, large quantities of data are required for the ongoing testing and refinement of algorithms. Misrepresentative and incomplete data can skew findings such that they are no longer applicable to the population that the AI was created for in the first place. While this can be addressed through data cleaning and building capacity to collect the vast amounts of health information needed to power AI, there is a growing concern of bias being embedded within the social systems we are analyzing (AI Now, 2016).

Second, regardless of how sophisticated our data collection systems are, if the underlying patterns reveal systemic biases, those will only be further embedded through deep learning algorithms (AI Now, 2016). For example, there are numerous studies on gender biases in the context of health care delivery. One study suggests that women in pain are more likely to be prescribed sedatives instead of pain medications (Calderone, 1990). If this health data were used to inform care then AI may learn from and replicate prescription patterns, further perpetuating this health care practice with effects disproportionately felt by the women in pain.

There is a need to reflect on the broader impact of data sharing on individual rights and liberties, including implications for individual autonomy and ownership of personal data (European Data Protection Supervisor Ethics Advisory Group [EDPSEAG], 2018; McKinsey Global Institute, 2017). There seems to be consensus on the need to reframe informed consent models within a data-driven economy such that citizens can be notified on how their information is being collected and what it is being used for (Centre for International Governance and Innovation, 2018). For instance, do citizens have the right to object to the sharing of their medical data to be used by AI? If practised widely, this could result in gaps in which population groups are represented, potentially compromising the integrity of AI decision making (Aspen Institute, 2017). As such, while within the scope of individuals' liberties, the right to object could potentially increase discrimination towards citizens who are underrepresented in such data (Aspen Institute, 2017).

Data Quality & Use

Questions were also raised on how data should be shared and who should be sharing it (McKinsey Global Institute, 2017). One such question arose around the tension between individual rights and the collective good. While designers and researchers should be obligated to ensure that they are collecting data from a wide variety of populations, what is an individual's civic responsibility to share their data for other populations to benefit? What public health functions would benefit from universal data sharing? How do we maintain individual autonomy through data sharing?

Another challenge is the commodification of data (European Data Protection Supervisor Ethics Advisory Group [EDPSEAG], 2018). What are the implications of attributing economic value to health data? The social consequences of trading such sensitive information may position people to either put their data on the market or risk being left out of the digital economy (EDPSEAG, 2018) and not reap the benefits that may come with utilizing AI in health.

Finally, lack of transparency about how algorithms operate (commonly know as the 'black box' problem) raises questions about the validity and trustworthiness of AI outputs (EDPSEAG, 2018). Collecting data for a predetermined purpose is difficult, and almost impossible, as the complexity and diversity of big data requires sophisticated algorithms to reveal underlying patterns. This creates a paradox in and of itself as the purpose of the data may not be known until an algorithmic analysis is performed (EDPSEAG, 2018).

Equity and the Digital Divide

The 'digital divide' (McKinsey Global Institute, 2017) describes the differing capacity of individuals, and health systems, to participate in the AI health ecosystem, potentially furthering health and social inequities.

Al technology can only be employed where there is broadband infrastructure and a sufficient quantity of data to support it. The current health care facilities pose limitations to data collection as electronic medical records (EMR) have yet to be adopted universally in not only Canada, but also globally (World Wide Web Foundation, 2017). Not only is the accuracy of the AI compromised due to insufficient data, but its applicability to diverse populations is also undermined (World Wide Web Foundation, 2017). As a result, AI systems may only produce outcomes for populations that are represented in the data, benefiting certain sectors of the population that have privileged access to said systems, and disadvantaging those that do not (AI Now, 2016).

Globally, where the infrastructure may not even yet exist, there is greater potential for a digital divide, whereby less wealthy and technologically advanced countries will be left disadvantaged in the global market, further exacerbating existing health and social inequities (International Telecommunication Union [ITU], 2017; United Nations Economic and Social Commission for Asia and the Pacific [UNESCAP], 2017). With an upfront disadvantage in data and computational resources, this inability to participate in, and draw benefits from the application of AI in health systems, may perpetuate a cycle where AI systems remain in the control of more powerful and wealthy nations (UNESCAP 2017). With an upfront disadvantage in data and computational resources, this inability to participate in, and draw benefits from the application of AI in health systems remain in the control of more powerful and wealthy nations (UNESCAP 2017). With an upfront disadvantage in data and computational resources, this inability to participate a cycle where AI systems remain in the control of more powerful and wealthy nations (UNESCAP 2017). With an upfront disadvantage in data and computational resources, this inability to participate a cycle where AI systems remain in the control of MI in health systems, may perpetuate a cycle where AI systems remain in the control of MI in health systems remain in the control of MI in health systems, may perpetuate a cycle where AI systems remain in the control of more powerful and wealthy nations (UNESCAP 2017).

The capacity to participate in the AI revolution constrains the possibility of ensuring a democratic development process to serve a diversity of populations and their complex needs, for participating in AI research is limited to those with the ability to pay (AI Now, 2016). Considering, for instance, that the majority of the private companies leading sector growth are situated in the Global North, there is significant concern of a "brain drain" (p. 10), as better pay and increased opportunities in high-income countries act as a significant pull factor for talent from other parts of the world (World Wide Web Foundation, 2017). This can further erode the capacity of LMICs to participate in and benefit from AI innovation in health.

7

Governance

The application of AI raises issues of governance in all sectors, including that of health, such as how its development and application will be regulated, why, and by whom. Eliciting public trust in AI through governance processes was one particular area identified as imperative for its successful implementation. One such aspect of this was health providers' trust in the evidence for AI as an effective health care delivery mechanism (The Center for Internet and Society India, 2017; The Standing Senate Committee on Social Affairs, Science and Technology, 2017).

Public trust concerns have also been raised around the responsible use and commercialization of data (Centre for International Governance and Innovation, 2018). Certain national jurisdictions, for instance, have significantly lower privacy standards for personal citizen information than others (Centre for International Governance and Innovation, 2018), which is something to consider in terms of international trade and what data, and potentially also algorithms, are subsequently purposed for. To address such issues, calls have been made to not only reform regulatory frameworks, but also to create a broader cultural shift and ongoing public dialogue that includes a social contract between citizens (data providers), and industry and government (data users) (British Academy for the Humanities and Social Sciences, and the Royal Society [BAHSS], 2017; The Centre for International Governance and Innovation, 2018).

Engaging in public dialogue will also ensure that regulatory policies are implemented consistently, rather than having Al's development and application operate within a patchwork of regulatory and governance structures, including across public and commercial sectors. Critics of a 'soft governance' (p.5) approach to AI, for instance, whereby AI development and application is governed by ethical codes of conduct, note the challenge of having to navigate numerous ethical codes with uncertainty regarding who will hold the responsibility of enforcement (AI Now, 2017). This is particularly pertinent to how and with whom a person, organization, or nation-state can seek redress for the negative consequences of AI, given the opaque nature of AI systems. Furthermore, without consistency, the global governance of AI is threatened by ethical loopholes, as was identified by the European Group on Ethics in Science and New Technologies (2018). In the absence of a coherent governance framework, an opportunity to 'ethics shop' (p. 14) presents itself, whereby AI developers and users can take advantage of jurisdictions with lower ethical standards (European Group on Ethics in Science and New Technologies, 2018).

Governance

This regulatory imbalance may similarly affect the marketing of AI products. If, however, innovators are unable to receive adequate funding for transforming their developments into products to take to market, they will look to market their products elsewhere (The Standing Senate Committee on Social Affairs, Science and Technology, 2017), where ethical standards of AI application are less robust. This speaks to the importance of consistent regulatory policies, but also to how best achieve the balance between obtaining government support for the commercialization of innovation (The Standing Senate Committee on Social Affairs, Science and Technology, 2017) and the checks and balances that may accompany such support, and pursuing what the Mercatus Center (2017) refers to as 'permissionless innovation'.

The 'permissionless innovation' (p.3) approach to AI governance relies on public policies not being guided by the precautionary principle (The Mercatus Center, 2017). Rather than pressing for transparency of innovation, they argue, AI developers should instead be able to improve their products without regulatory hindrance in order to better meet the needs of different populations, including those who are marginalized (The Mercatus Center, 2017). An opposing argument is against this proprietary approach, and is instead in favour of availing AI algorithms for open-sourced and public use, to enhance the benefit of AI for all. Given these concerns for responsible innovation, among others, the British Academy for the Humanities and Social Sciences (BAHSS) (2017) has called for a new governance framework, based on transcending principles to suit all sectors to which AI is applied, and which include both voluntary standards and regulation. One such principle includes ensuring that decisions made for the management and use of data are done so in way that is transparent, accountable, and inclusive, with the overarching principle of promoting human flourishing (BAHSS, 2017). They also recommend establishing a new independent body to monitor the entirety of the AI governance landscape, to ensure these transcending principles are being utilized within both the public and private sectors, to ensure innovation and commercial interest is balanced against public good (BAHSS, 2017).

Collaboration

Progress within the field of AI requires meaningful collaboration and inclusion of various stakeholders from the design and development of AI technologies to their application and governance. Currently, it has become clear that the private tech industry is leading AI innovation (AI for Good, 2017). There has also been an increase in private sector investment in ethics infrastructure both within technology organizations and within civil society through philanthropy donations (AI for Good, 2017). Collaborations like the The Partnership on AI, which was founded by researchers at Amazon, Google, IBM, Microsoft, Facebook, and Apple, are aiming to create open platforms to discuss and identifies opportunities to advance the social purposes of AI (AI for Good, 2017). Interestingly, industry representatives themselves have shared their concern for steering innovation towards profit driven applications of AI rather than those with high social value (AI for Good, 2017).

Initiatives such as a Global Fund for AI for Social Good (AI for Good, 2017) would aim to prioritize and address the world's most pressing health and social challenges by providing a platform for investments in AI research and applications that focus on the Sustainable Development Goals for the most vulnerable populations (AI for Good, 2017). A recurring theme in the literature is that countries need to collectively identify areas for partnerships and create international frameworks that will promote AI research and development (National Science and Technology Council Committee on Technology, 2016). To ensure that AI is responsive to the unique needs of nation-states, it is imperative that global decision making reflect the ideas and realities of nations across the globe, especially those from low and middle income countries (World Wide Web Foundation, 2017).

The need for collaboration extends beyond global innovation and governance of AI to the very people that are ideating and developing AI, with experts calling for a participatory approach to AI development (British Academy for the Humanities and Social Sciences, and the Royal Society, 2017). For example, including users of technology, such as clinicians, into the design process ensures that AI is applicable and representative of the needs of the users that will ultimately utilize the technology (AI Now, 2017). Accordingly, interdisciplinary collaboration with experts beyond the AI field will ensure AI is responsive to societal needs (AI Now, 2017). AI developers should be encouraged to work alongside ethicists and other social scientists who are well-versed in ethical and social implications of AI use to ensure technologies are developed with the necessary foresight of its impact on humanity (AI Now, 2017).

Collaboration

A notable trend in the grey literature scanned thus far, is that with collaborations emerging between academics and the private sector, many authors now wear "multiple hats" and represent numerous organizations and sectors in their work. For example, a commentary in an academic journal can be authored by an individual working at an academic institution who is also the Chief Scientific Advisor at a private tech company. Under whose authority, and from which organizational perspective are policy documents being written? Who is ultimately driving the research agenda? What, if any, are the implications of researchers who are wearing "multiple hats" for the reliability, accountability, and transparency of research in AI?



Redefining Health Care Practice

Health care is a primary focus of AI development and innovation across a range of areas, including disease prevention, diagnosis, and treatment. It has been described as a high growth area (McKinsey Global Institute, 2017). What impact will AI-mediated care have AI on the patient-clinician relationship and on clinician and patient roles in the clinical setting?

While some suggest the incorporation of AI in health care practice will give physicians more time to engage in empathetic discussions with their patients, others are concerned that AI could reduce the care setting to a mere technical encounter (Nuffield Council of Bioethics, 2018). Al is primarily trained on quantitative measures such as images, health records and blood tests, which runs the risk of reducing clinical diagnosis to only measurable data and failing to capture the nuances of a patient's values or social, economic, and cultural situation and the verbal and nonverbal cues that inform clinical judgment about a particular patient (Future Advocacy, 2018). Furthermore, if AI is fully integrated into care settings, will patients have the capacity to refuse care from AI? Is a patient's right to health synonymous with a right to access health care delivered by clinicians (Future Advocacy, 2018)? And as computational algorithms become increasingly sophisticated and reliable, what role will or should clinical judgment play? Could the diffusion of AI through clinical processes have the effect of displacing clinical experience as the gold standard of knowledge in health (Char, Shah, Nigam, and Magnus, 2018)?

In response to concerns about preserving the relational core of the clinician-patient interaction, it has been argued that the main use of AI should be to complement clinician judgement not to replace it, and that the development of AI in health care should be undertaken with the clinician-patient relationship in mind (The Centre for Internet and Society, India, 2016). However, this leaves as yet unanswered how discrepancies between AI-mediated outputs and clinical judgment ought to be resolved. Clinician trust of AI outputs is critical if AI-mediated care is to realize its benefits for patients. Lack of transparency about how AI systems work, particularly how machine learning generates a clinically-relevant output, is a particular concern (AI Now, 2017). In addition, if a medical error occurs as a result of applying an AI-mediated diagnostic or treatment decision, who should hold accountability for the error - the AI system, the developers of the technology, or the clinician delivering the care?

Redefining Health Care Practice

Finally, the proliferation of health-related digital apps and devices are anticipated to redefine the role that patients play in their health care planning (AI Now, 2016). These technologies may empower patients to make informed choices about their health. On the other hand, being informed means being able to access, understand, and apply the information to their particular health needs, which entails a certain level of health literacy, and digital proficiency as well as a receptive health care system. Where these technologies operate independently of health care systems, users may have difficulty interpreting the information and taking appropriate action. Concerns have also been raised about whether some direct-to-consumer health information technologies may place an additional burden on health care systems and increased health care utilization in response to consumer demand (RAND, 2017). Finally, questions have been raised about whether the wider use of patient or consumer-level digital apps in care of individual patients may advantage those with a high level of health literacy, time and resources and disadvantage those who do not (AI Now, 2016).



Future of Work

Al has the potential to disrupt work within the health sector. The impact of this disruption is not well-understood; however, its effect on health and social inequity is a central concern, particularly for women and marginalized groups. Increasing gender wage gaps have been flagged as a potential outcome of Al-mediated work (McKinsey Global Institute, 2017; World Wide Web Foundation, 2017). First, this may result from the automatization of certain types of work involving routine tasks, which are often low-paid and over represented by women (McKinsey Global Institute, 2017; World Wide Web Foundation, 2017). It may also be an unintended consequence of economic investment in STEM (Science, Technology, Engineering, and Math) jobs, which are well-paid and male-dominated (World Wide Web Foundation, 2017). Some argue that new employment opportunities will emerge as we have observed with workforce disruptions throughout history; however this fear of displacement remains for any worker holding a job with the potential to be automated, jobs which are typically held by already low-paid workers (Mercatus Center, 2017).

The application of AI in health care has significant implications for the training and reskilling of health workers. Assuming health workers deliver care with the support of AI, it is imperative that they be equipped with the digital skills necessary to deploy these technologies (Mesko, 2017). By understanding how AI works, health workers will be better able to assess and apply AI outputs in clinical decisionmaking (The Centre for Internet and Society India, 2017). This is relevant not only for patient well-being, but also in terms of the displacement of health workers in particular. While it is unlikely that AI will replace, for example, physicians, physicians able to apply AI in practice may replace those who do not (Mesko, 2017). Efforts to educate health workers will need to integrate digital skills training into current curricula to ensure that the future workforce is well prepared to deliver care alongside AI. It is also necessary to integrate social sciences and humanities into the training curriculum for future developers and programmers (World Wide Web Foundation, 2017). By being vigilant about concerns of bias, privacy, and consent, those involved in the development of AI may be able to prevent or mitigate any potential negative consequences (World Wide Web Foundation, 2017).

Future of Work

AI technologies hold the potential to enhance the operational efficiency of health systems with implications for the management of human resources. Through ongoing surveillance and data collection, AI can be utilized to monitor employees, potentially without their knowing. It can also be used in scheduling by predicting when workers are most needed (AI Now, 2017). This under-explored area of risk may further contribute to an ever-increasing shift towards 'gig work' (AI Now, 2017, p. 11), including in the health field, particularly in roles that are not patient-facing. While unpredictable schedules would mean increased profit margins for governments and companies managing the sector, workers themselves may be negatively affected by the uncertainty of a precarious income and other stressors that come with gig work (AI Now, 2017). AI may also impact hiring practices in the workforce more broadly. If the AI system has been trained by information it receives from past hiring decisions, decisions which may have been biased, then the information utilized to screen new applicants may reflect said bias (National Science and Technology Council Committee on Technology, 2016), and disadvantage certain population sub-groups from entering the workforce.

Conclusion

Through this preliminary scan of the grey literature, we have identified numerous themes and challenges that are emerging in the field of AI, health, and ethics in the context of a globalized world. Though these issues were thematically organized, we must emphasize how interconnected each of these challenges are. For example, when considering the need for greater collaboration to incentivize ethically and socially minded AI innovation, we cannot do so without the appropriate governance to regulate such technologies. Similarly, as health care practice is redefined by AI and integrated into clinical care, there is a growing need to educate and re-skill the health care workforce. A notable finding of work thus far is a predominant focus on the health care context in discussions of AI for health. The implications of AI for population and public health is a critical gap that needs to be addressed. The views of patients and the public on AI for health are also not well-understood and present additional opportunities for civic engagement, research and education. While we have only scratched the surface, we hope that the review and analysis provided in this backgrounder has further inspired questions and raised points of discussions to be addressed in the coming days.

We thank you for joining us for the inaugural symposium on Ethics and AI for Good Health. For more information about our work and to explore opportunities for collaboration, contact us at jcb.director@utoronto.ca.

References

Barton, D., Woetzel, J., Seong, J., & Tian, Q. (2017). *Artificial intelligence: Implications for China.* McKinsey Global Institute. Retrieved from https://www.mckinsey.com/featuredinsights/china/artificial-intelligenceimplications-for-china

Bollier, D. (2017). Artificial intelligence comes of age: The promise and challenge of integrating AI into cars. healthcare and journalism: A report on the inaugural Aspen Institute roundtable on artificial intelligence. Aspen Institute. Retrieved from http://csreports.aspeninstitute.org/documents/AI2016Rpt.pdf

British Academy for the Humanities and Social Sciences & the Royal Society. (2017). *Data management and use: Governance in the 21st century.* Retrieved from https://royalsociety.org/~/media/policy/projects/data-governance/datamanagement-governance.pdf

Burgess, J., Floridi, L., Pols, A., & van den Hoven, J. (2018). *Towards a digital ethics.* European Data Protection Supervisor Ethics Advisory Group. Retrieved from https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf

Calderone, K. L. (1990). The influence of gender on the frequency of pain and sedative medication administered to postoperative patients. *Sex Roles*, 23(11-12), 713-725.

Centre for International Governance and Innovation. (2018). *A national data strategy for Canada: Key elements and policy considerations*. Retrieved from https://www.cigionline.org/sites/default/files/documents/Paper%20no.160.pdf

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care — addressing ethical challenges. *New England Journal of Medicine*, 378(11), 981-983. doi:10.1056/nejmp1714229

Compolo, A., Sanfilippo, M., Whittaker, M., & Crawford, K. (2017). *AI Now 2017 report.* AI Now Institute. Retrieved from https://ainowinstitute.org/AI_Now_2017_Report.pdf

Corbett, J., d'Angelo, C., Gangitano, L., & Freeman, J. (2017). Future of health: Findings from a survey of stakeholders on the future of health and healthcare in England. RAND Corporation. Retrieved from https://www.rand.org/content/dam/rand/pubs/research_reports/RR2100/RR2147/R AND_RR2147.pdf

References

Crawford, K. & Whittaker, M. (2016). *The AI Now report: The social and economic implications of artificial intelligence technologies in the near-term.* AI Now Institute. Retrieved from https://ainowinstitute.org/AI_Now_2016_Report.pdf

European Group on Ethics in Science and New Technologies. (2018). *Statement on artificial intelligence, robotics and autonomous systems*. Retrieved from http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

Executive Office of the President National Science and Technology Council Committee on Technology. (2016). *Preparing for the future of artificial intelligence*. Retrieved from

https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsit es/ostp/NSTC/preparing_for_the_future_of_ai.pdf

Fenech, M., Strukelj, N., & Buston, O. (2018). *Ethical, social, and political challenges of artificial intelligence in health.* Future Advocacy. Retrieved from https://wellcome.ac.uk/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf

Gerstein Library (2018). Searching the literature: A guide to comprehensive searching in the health sciences. Retrieved from https://guides.library.utoronto.ca/c.php?g=577919&p=4123572

International Telecommunication Union. (2017). *AI for good global summit report.* Retrieved from https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf

Mesko, B. (2017). The role of artificial intelligence in precision medicine. *Expert Review of Precision Medicine and Drug Development*, 2(5), 239-241. doi:10.1080/23808993.2017.1380516

Mohandas, S. (2017). *AI and healthcare in India: Looking forward roundtable report.* The Centre for Internet and Society, India. Retrieved from https://cisindia.org/internet-governance/files/ai-and-healthcare-report

Nuffield Council on Bioethics. (2018). *Artificial intelligence (AI) in healthcare and research.* Retrieved from http://nuffieldbioethics.org/wp-content/uploads/Artificial-Intelligence-AI-in-healthcare-and-research.pdf

References

Standing Senate Committee on Social Affairs, Science and Technology. (2017). Challenge ahead: Integrating robotics, artificial intelligence and 3D printing technologies into Canada's healthcare systems. Retrieved from https://sencanada.ca/content/sen/committee/421/SOCI/reports/RoboticsAI3DFinal _Web_e.pdf

Thierer, A., O'Sullivan, A., & Russel, R. (2017). *Artificial intelligence and public policy.* The Mercatus Center. Retrieved from https://www.mercatus.org/system/files/thierer-artificialintelligence-policy-mr-mercatus-v1.pdf

United Nations Economic and Social Commission for Asia and the Pacific. (2017). *Artificial intelligence in Asia and the Pacific.* Retrieved from http://www.unescap.org/sites/default/files/ESCAP_Artificial_Intelligence.pdf

World Wide Web Foundation. (2017). *Artificial intelligence: The road ahead in low and middle-income countries.* Retrieved from https://webfoundation.org/docs/2017/07/AI_Report_WF.pdf