THE ROLE OF OPEN SOURCE AND OPEN DATA IN THE REGULATION OF MEDICAL ARTIFICIAL INTELLIGENCE

Submission to the Federation of Medical Regulatory Authorities of Canada

Puneet Kapur MD, FRCPC & Glen Vajcner MD, FRCSC

The Role of Open Source and Open Data in the Regulation of Medical Artificial Intelligence

"When you put together open medicine, open science, open access, open source, and open data—Open5—all sorts of new channels of research activity become available, and existing ones become exponentially more powerful."

— Eric Topol from "The Patient Will See You Now: The Future of Medicine is in Your Hands"[1]

"Alchemists turned into chemists when they stopped keeping secrets."[2]

~ Eric S. Raymond. Celebrated Computer Scientist and founder of the "Open Source Initiative"

Introduction

Regulation of the medical profession in Canada is divided along clear functional lines. Pharmaceuticals and medical devices are regulated by Health Canada, the Medical Council of Canada assesses educational competency, and finally the provincial and medical regulatory authorities (MRA) govern the clinical practice of medicine. At first glance, the divisions seem sensible with different requirements and rules for technology used by physicians, aspiring physicians and practicing physicians making decisions at the bedside. Yet, when considering the impact of artificial intelligence (AI) on modern medicine, the divisions become blurred.

When faced with a technology that is capable of self-guided unstructured learning[3] that can integrate knowledge to make diagnoses[4], where does the responsibility for regulation lie? It would be overly simplistic to attempt regulation of AI as though they were simple medical devices like ECG machines or ultrasound machines. Thus, the stated goal on the FMRAC website that since "the MRAs are regulators of physicians (and not of technologies or health systems), [and essay] submissions must fall within this scope" is misleading. We are faced with a new era of 'thinking machines' and this requires a new approach to legislation that transcends the current regulatory framework. We argue that while MRAs are 'not [regulators] of technologies' they must become intimately familiar with technology and comment on the types of technology and how it is used for the best interests of patients. We propose an argument that MRAs role is to strive to advise physicians and develop rules around emerging technologies, while ensuring the foundation on which these technologies are built are free of biases that may lead to inappropriate practice patterns supported by these technologies. MRAs may facilitate this by supporting two core technical principles; Open Source and Open Data. In turn each of these principles has profound implications on the areas of interest identified by FMRAC namely, data protection, ethical obligations to society and our responsibility to patients.

Open Source and Data Protection

Most physicians are personally familiar with commercial software wherein "users pay for a software license with strict restrictions and no access to the source code". In contrast 'Open Source' or 'Open-Source Software' (OSS) refers specifically to software where " end users do not necessarily have to pay, the license has fewer restrictions and the software includes the source code which they can examine,

modify and incorporate into their own system"[5]. At first glance the distinction between not being able to access source code (i.e., commercial software) and full access to view and modify the source code of medical software (i.e., open-source software) would seem be a technical detail well outside the purview of medical regulatory authorities, yet it has wide reaching repercussions.

Protecting the personal health information (PHI) of patients is required by the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) and other supplementary provincial privacy legislations. It would seem obvious that the best way to keep electronic PHI confidential is by using software where there are similarly 'strict restrictions and no access to the source code'[5]. This idea has been aptly termed 'security through obscurity' and has been repeatedly shown to offer no real security at all. As one author noted an "example of security through obscurity is hiding the key to your front door under the doormat. If the burglar knows the location of the key, the scheme is not secure anymore'[6]. In fact, the world's most secure systems rely on Kerckhoffs's principle which states that "the security of a system should not rely on keeping the details of the devices or algorithms private"[7]. We are not the first to make the argument that open source is safe and secure for use in healthcare[8–10] but we are the among the first to argue that Canadian medical regulatory authorities should encourage physicians, whenever possible, to consider open source software tools over closed source alternatives, because the latter can provide intrinsically greater data protection. With a developed in-house expertise, an MRA may analyze the building blocks of the tools that physicians in their jurisdiction will use to treat patients and simultaneously ensure that these blocks provide robust protection of PHI. In doing so we join efforts by the regulators in India, Brazil, Caribbean States and the Pan American Health Organization who have endorsed national adoption of OSS[11]

Open Data and Ethical Responsibilities

The term artificial intelligence conveys a sense of rationality and impartiality, but a considerable body of work has shown that AI decision making can be biased against women, minorities, and people of color[12–15]. Human decision making becomes biased in subtle ways based on our training and machines are remarkably similar. Artificial intelligence relies on huge volumes of 'training data' to learn from and skewed training data gives raise to AI biases. In a notable example, Amazon[16] abandoned an artificial intelligence recruiting tool that showed a bias against women. Since the tool had been trained on the last 10 years of Amazon hiring data, it adopted the biases hidden within that data. Other work has argued for increased representation of people of color in AI models used to detect skin cancer[17,18].

One solution to the problem of bias in AI is the use of open data, in particular open data training sets. To ensure equity and equality, organizations that creates AI should release the datasets used to train the AI tools so it can be scrutinized for bias. For example, Google has released a 10 million image dataset to improve machine vison. Despite the advantages of open datasets, the companies and organizations that create AI for medical applications are not mandated to use known bias-free datasets to train their AIs nor are they required to release their datasets for public scrutiny. MRAs ensure that physicians meet academic standards and pursue ongoing medical education, but as AI tools become more wildly used, MRAs will need to ensure that they too have an appropriate 'education'. We have seen that the regulatory framework governing potentially biased data has lagged behind the development of AI[19]. A recent retrospective analysis of one commercial deployed AI showed that it under-estimated severity of illness in African Americans[20]. Using the influence of medical regulatory authorities to establish fairness

in access to healthcare is well established and now some jurisdictions such as New York are extending their authority to make "algorithms accountable"[21].

Enforcement of detailed technical standards for AI data training sets is beyond the scope of regulatory authorities, however arguing for principles that lead to the creation of bias-free training datasets is not. The New York City committee on technology drafted a bill to address 'Automated Decision Systems Used By Agencies' and argued that "whenever a city agency wished to use an automated system to apportion policing, penalties, or services, the agency would … be required to simulate the algorithm's real-world performance using data submitted by New Yorkers"[21]. Canadian medical regulatory authorities should adopt similar ethical principles and encourage physicians to be aware biases in the AI tools they employ.

Open Source, Open Data and the Patient Protection

A central mandate of MRAs is to prevent patient harm and adjudicate cases when harm does occur. The argument is often made that in cases of adverse events, the physician leading the healthcare team bears most of the responsibility. If a surgical resident accidentally cuts an artery, or a radiologist fellow misses a cancerous growth, it is the attending staff member who, as the 'most responsible physician', shares the blame. However, who is responsible if the 'most responsible physician' is not a physician at all? If an autonomous surgical robot cuts an artery or a radiology algorithm operating independently, misses a tumor; then where should responsibility be placed?

The ethical dilemma of decisions made by thinking machines is not unique to medicine. The United Nations (UN) has discussed moral responsibility when using autonomous weapons[22], and more recently the National Transport Safety Board (NTSB) in the United States provided guidance on liability when fatalities are caused by self driving vehicles[23]. Neither the United Nations nor the NTSB have reached conclusions regarding the rules around artificial intelligence, but their efforts are instructive. In trying to establish the source (and by extension responsibility) for autonomous vehicle fatalities, the NSTB included in its post crash analysis, a detailed analysis of the underlying software[24] and made recommendations to repair underlying software flaws[25]. Similarly, the United Nations Institute for Disarmament Research (UNDIR) has argued that AI technology is often dual use[26] and others have argued the software used to keep pedestrians safe in civilian autonomous vehicles could be used to target pedestrians in military vehicles. Accordingly the UNDIR also calls for access to the underlying software to ensure "transparency, interpretability, validation and some form of verification for such systems"[26].

In looking at work of NTSB, UNDIR and others striving to create a regulatory environment for AI, the common theme is that assignment of responsibility to man or machine requires access to the detailed decision-making process of both. Just as MRAs currently request physician documentation in investigations of where patients were harmed, they must also be able to request details of both the source code and training data used in instances where an artificial intelligence causes harm. Thus, the argument for open-source AI built on open data comes full circle.

Summary

Medical regulatory authorities have not typically been in the business of regulating technology because, until now, medical devices were not capable of fully independent action. Some authors[27,28] still argue that AI is merely a new type of medical device and can eventually be folded into the current regulatory regime. We argue that the emergence of machines that can learn and make their own decisions is a paradigm shift that calls upon MRAs to take a more hands approach to regulated technology.

FMRAC is a keystone institute in Canadian healthcare. It is at this level that larger philosophical questions of transparency, equality and openness of the health care system should be debated. Position statements from the FMRAC set the 'moral tone' that shape discussions in provincial colleges of physicians and surgeons and research institutions around the country. Taking a position on advocating for the use of open source and open data will help ensure data protection, bias free algorithms, and ensure patient safety. Commenting on New York City's attempt to hold AI accountable, Frank Pasquale, a law professor from the University of Maryland, summed up the need for openness well when he stated, "Secrecy may incentivize tiny gains in efficiency, but those are not worth the erosion of legitimacy and public confidence in government. It's a dereliction of duty to allow vital decisions to be made by a black box"[21].

REFERENCES

1. Topol E. *The Patient Will See You Now: The Future of Medicine Is in Your Hands*. Basic Books; 2016.

2. Raymond ES. *The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'Reilly Media; 2001.

3. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4). doi:10.1136/svn-2017-000101

4. Ahmed MN, Toor AS, O'Neil K, Friedland D. Cognitive Computing and the Future of Health Care Cognitive Computing and the Future of Healthcare: The Cognitive Power of IBM Watson Has the Potential to Transform Global Personalized Medicine. *IEEE Pulse*. 2017;8(3):4-9. doi:10.1109/MPUL.2017.2678098

5. McDonald CJ, Schadow G, Barnes M, et al. Open Source software in medical informatics—why, how and what. *Int J Med Inf*. 2003;69(2):175-184. doi:10.1016/S1386-5056(02)00104-1

6. Lydersen LVV de W. Practical security of quantum cryptography. Tapir Akademisk Forlag. 2011. Published online 2011.

7. Singh S. *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Anchor; 2000.

8. Reynolds CJ, Wyatt JC. Open Source, Open Standards, and Health Care Information Systems. *J Med Internet Res*. 2011;13(1):e24. doi:10.2196/jmir.1521

9. Webster PC. Canada's ehealth software "Tower of Babel." *CMAJ*. 2010;182(18):1945-1946. doi:10.1503/cmaj.109-3717

10. Goulde M, Holt M. *Open Source Software: A Primer for Health Care Leaders*. California HealthCare Foundation; 2006.

11. Aminpour F, Sadoughi F, Ahamdi M. Utilization of open source electronic health record around the world: A systematic review. *J Res Med Sci Off J Isfahan Univ Med Sci*. 2014;19(1):57-64.

12. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. *J Glob Health*. 9(2). doi:10.7189/jogh.09.020318

13. Buolamwini J, Gebru T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: *Conference on Fairness, Accountability and Transparency*. PMLR; 2018:77-91. Accessed February 2, 2021. http://proceedings.mlr.press/v81/buolamwini18a.html

14. Gebru T. Race and Gender. The Oxford Handbook of Ethics of AI. doi:10.1093/oxfordhb/9780190067397.013.16

15. Treating health disparities with artificial intelligence | Nature Medicine. Accessed February 2, 2021. https://www.nature.com/articles/s41591-019-0649-2

16. Amazon scraps secret AI recruiting tool that showed bias against women - Reuters. Accessed February 2, 2021. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

17. Adamson AS, Smith A. Machine Learning and Health Care Disparities in Dermatology. *JAMA Dermatol*. 2018;154(11):1247-1248. doi:10.1001/jamadermatol.2018.2348

18. AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind - The Atlantic. Accessed February 2, 2021. https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/

19. Pesapane F, Volonté C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights Imaging*. 2018;9(5):745-753. doi:10.1007/s13244-018-0645-y

20. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. doi:10.1126/science.aax2342

21. New York City's Bold, Flawed Attempt to Make Algorithms Accountable | The New Yorker. Accessed February 2, 2021. https://www.newyorker.com/tech/annals-of-technology/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable

22. Warren A, Hillas A. Lethal Autonomous Weapons Systems: Adapting to the Future Unmanned Warfare and Unaccountable Robots. *Yale J Intl Aff*. 2017;12:71.

23. Kelley B. Public health, autonomous automobiles, and the rush to market. *J Public Health Policy*. 2017;38(2):167-184. doi:10.1057/s41271-016-0060-x

24. Plaza L. Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian, Tempe, Arizona, March 18, 2018. :78.

25. Shepardson D. In review of fatal Arizona crash, U.S. agency says Uber software had flaws. *Reuters*. https://www.reuters.com/article/us-uber-crash-idUSKBN1XF2HA. Published November 6, 2019. Accessed February 3, 2021.

26. *The Weaponization of Increasingly Autonomous Technologies: Autonomous Weapon Systems and Cyber Operations*. United Nations Institute for Disarmament Research; 2017. Accessed January 30, 2021. https://unidir.org/files/publications/pdfs/autonomous-weapon-systems-and-cyber-operations-en-690.pdf

27. Harvey HB, Gowda V. How the FDA Regulates AI. *Acad Radiol*. 2020;27(1):58-61. doi:10.1016/j.acra.2019.09.017

28. Food US, Administration D. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)-Discussion Paper and Request for Feedback. Published online 2019.