



CIFAR

MACHINE MD:

Law and Ethics of Health-Related AI Case Study 4:

Cardiac Arrest Prediction

Workshop held: October 14, 2022

Report published: March 10, 2023

This report was drafted by Sophie Nunnelley in collaboration with the participants of the Machine MD: Law and Ethics of Health-Related AI Case Study 4: Cardiac Arrest Prediction.

Acknowledgments

This event was co-hosted by CIFAR and the Canadian Institutes of Health Research (CIHR)-funded Machine MD: How Should We Regulate AI in Health Care? project, with support from the Alex Trebek Forum for Dialogue. It is part of CIFAR's AI & Society Program. CIFAR's leadership of the Pan-Canadian AI Strategy is funded by the Government of Canada, with support from Facebook and the RBC Foundation. The organizers thank CIFAR, CIHR, and the Alex Trebek Forum for Dialogue for their support.

Citation

S. Nunnelley, A. Goldenberg, C. Régis, C. M. Flood, T. Scassa, A. Ferron Parayre, P. Déziel, V. Gruben and the workshop participants, Machine MD: Law and Ethics of Health-Related A.I. Case Study 4: Cardiac Arrest Prediction (Toronto: CIFAR, 2023).

Table of Contents

Law and Ethics Case Studies in Health-Related AI	4
AI-Based Cardiac Arrest Prediction in the Pediatric ICU (Dr. Anna Goldenberg, SickKids)	5
Commentaries	7
A. Informed Consent (Audrey Ferron Parayre, University of Ottawa Faculty of Law)	7
B. Privacy (Pierre-Luc Déziel, Université de Laval, Faculté de droit)	10
C. Liability (Vanessa Gruben, University of Ottawa, Faculty of Law)	12
Breakout Sessions	14
Breakout #1: Informed Consent	14
Breakout #2: Privacy	16
Breakout #3: Liability	17
DISCUSSION	18
I. Interplay between the different issues	18
II. The challenge of informed consent	18
III. The law - practice gap and the importance of patient perspectives	20
Conclusion	22

Law and Ethics Case Studies in Health-Related AI


The CIHR-funded Machine MD: How Should We Regulate AI in Health Care? project is led by Anna Goldenberg (Senior Scientist, SickKids); Catherine Régis (Law, Université de Montréal) Colleen M Flood (Law, University of Ottawa); and Teresa Scassa (Law, University of Ottawa). The project is dedicated to investigating the legal and ethical issues raised by artificial intelligence (AI) in health care and to developing recommendations for their optimal governance. Part of the Machine MD team's work includes examining real AI technologies, the practical issues they raise, and their current treatment in Canadian and foreign law. This approach moves beyond abstract concerns into concrete realities, helping to inform law reform with a better understanding of real-world applications. The goal is to support beneficial AI technology innovation, while minimizing associated risks through appropriate legal governance.

In keeping with this aim, the Machine MD team has partnered with CIFAR to host a series of online case study events.¹ Each event assembles an interdisciplinary group of experts in AI, law, ethics, policy, and medicine to discuss the regulatory issues raised by a specific AI technology. The previous three case studies took place in the spring of 2022, and concerned the “OR Black Box”, the Suicide Artificial Intelligence Prediction Heuristic or “SAIPH”, and “digital twins” technology.² This report summarizes the findings of the fourth case study in the series, concerning the development of an AI-based cardiac arrest prediction technology. The event brought together 36 experts with a range of backgrounds and perspectives.

¹ See CIFAR, “AI & Society: Advancing our understanding of the ethical, legal, political, and society implications of AI”, online: <<https://cifar.ca/ai/ai-and-society/>>.

² Reports from the previous events are forthcoming and will be available online at <www.cifar.ca>.

AI-Based Cardiac Arrest Prediction in the Pediatric ICU (Dr. Anna Goldenberg, SickKids)



The event began with a presentation from Anna Goldenberg, Senior Scientist in the Genetics and Genome Biology program at SickKids Research Institute, and one of the developers of the cardiac arrest prediction technology.³ Melissa McCradden (SickKids) and Sana Tonekaboni (University of Toronto / Vector institute) also assisted in answering questions.

Goldenberg explained that the technology began with the goal of predicting cardiac arrest in the pediatric intensive care unit (ICU) at SickKids hospital in Toronto. The motivating example was that of a nine-year-old boy who fractured his femur while skiing. A clot formed during the boy's surgery, which led to cardiac arrest and brain damage. Goldenberg explained that this is a tragic but not unusual course of events; the statistics relating to cardiac arrest are not favourable, with the chance of survival following cardiac arrest being approximately 14% overall and 36% if the person is in hospital. Adverse side effects such as brain damage are, moreover, common. Goldenberg and the other members of the development team were motivated by the knowledge that *prevention* is a key to avoiding these numbers and outcomes.

A problem with creating a predictive model is, however, the sheer volume of data created in a typical ICU. Patients are connected to monitors and machines that measure heart rate, pulse, central venous pressure, blood pressure, airway respiratory rate, among other physiological signals, generating a volume of data that is comparable to the flow-through of Niagara Falls.⁴ Goldenberg explained that it is simply not possible for humans to detect all aberrations and patterns in this data, which frequently overwhelms care providers. This is where AI comes in; Goldenberg and her colleagues considered this context to be ideal for computational algorithms and machine learning. She explained that their model can take in all this patient data and, with little human intervention, assess the probability of cardiac arrest, thus allowing for earlier

³ The cardiac arrest prediction model was developed by a team that includes Peter Laussen (SickKids), Michael Brudno (SickKids), Sana Tonekaboni (University of Toronto/Vector Institute), Mjaye Mazwi (SickKids), Robert Greer (SickKids) and others.

⁴ Goldenberg attributed this comparison to Peter Laussen, who has reportedly compared the data that must be monitored in the ICU (200,000 bytes/sec) to the volume of water moving through Niagara Falls (200,000 ft³/sec).

intervention.

At same time, Goldenberg noted, the tool has shifted beyond cardiac arrest prediction, due in part to the problem of false positives. There are only about 100 cardiac arrests per year in the SickKids Critical Care unit. Assessing risk every 5 minutes, for 30 beds, over 365 days per year, would result in approximately 30,000 false positives a year. This could in turn cause alarm fatigue among clinicians, leading them to ignore the system, and fail to act in instances of actual risk.

Goldenberg explained that their solution was to expand the purpose of their tool – to include not just risk of cardiac arrest, but also, *risk of deterioration*. Their model takes the physiological symptoms that are collected at the bedside and uses them to provide a deterioration risk score that applies to a two-hour window. Physicians can then use that score to make decisions about treatments and interventions. Goldenberg emphasized that the tool does not collect new data; it summarizes existing data streams to assess risk. Moreover, the tool is imperfect and will make mistakes. However, she said the physicians working with them know this and accept that they are ultimately responsible for making treatment decisions.

Commentaries



Following Goldenberg's presentation the group heard from legal experts on three pre-identified issues – informed consent, privacy, and liability – that frequently arise in the context of health-related AI. These presentations were intended to provide an overview of these areas of law and their potential applications (along with gaps and ambiguities) to this prediction technology.

A. Informed Consent (Audrey Ferron Parayre, University of Ottawa Faculty of Law)

Audrey Ferron Parayre began the legal presentations with a discussion of the law of informed consent. She explained that the applicable consent standards will vary with the context, including, whether we are concerned with (1) consent to use the tool at the point of care for research purposes (implementation research); (2) consent to employ the tool in clinical care; or (3) consent in the context of clinical care for minors.

On the subject of research, Ferron Parayre explained that research involving patients requires explicit consent, unless the research will not adversely affect the care that is received; it poses no more than a minimal risk to the person; and obtaining consent would adversely affect the research.⁵ She opined that in the case of the cardiac arrest prediction technology, consent should be obtained prior to any implementation research, given that the tool will influence the person's care, and there is no indication that obtaining consent would adversely affect the research. Moreover, she explained the consent standard is higher here than that it is in the context of consent to care.⁶

Turning to informed consent to healthcare uses of the technology, Ferron Parayre began by emphasizing the underlying autonomy principles: Informed consent is intended to protect dignity and autonomy. Moreover, it helps to build trust between the patient and clinician, a consideration that is perhaps especially important to the implementation of novel AI technologies in healthcare settings. The question, however, is what these principles ought to mean in this context. Indeed, Ferron Parayre raised two core questions: Under current law, must patients (or their substitute

⁵ Government of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2022)*, Chapter 3, The Consent Process, online: <https://ethics.gc.ca/eng/policy-politique_tcps2-epc2_2022.html>.

⁶ *Halushka v University of Saskatchewan* (1965), 53 DLR (2d) 436 (Sk CA); 1965 CanLII 439.

decision-makers) give informed consent before they can be monitored by an AI algorithm in the ICU? And second, should the law require such consent?

Ferron Parayre discussed some Canadian work on these questions, noting that it has led to conflicting recommendations and approaches. The Royal College of Physicians and Surgeons of Canada's Task Force Report on Artificial Intelligence and Emerging Technologies discusses the need for patient consent to the collection and use of data but does not speak to whether use of an AI algorithm in healthcare requires informed consent.⁷ Quebec legislation and the Office of the Privacy Commissioner of Canada (OPCC), on the other hand, have both addressed the matter of informed consent to AI in healthcare, but have come to different conclusions. Article 65.2 of Quebec's *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels* only mandates disclosure that a decision is based in AI if that decision is made without any human input.⁸ As such, it would not require that patients consent to the use of this cardiac arrest predictor tool, as that technology keeps a physician "in the loop". On the other hand, the OPCC's report - *A Regulatory Framework for AI: Recommendations for PIPEDA Reform* – suggests the obligation to disclose the use of AI should not turn on whether a human is involved in the decision.⁹ Ferron Parayre suggested that the Quebec law should have been more aligned with this latter recommendation.

The key question in determining whether there must be disclosure and informed consent to use this prediction technology in healthcare, Ferron Parayre explained, is whether this is material information. While the statement of this principle is different in Quebec civil law (which asks what a "reasonable physician" would disclose) and in common law provinces (where the question is what a "reasonable patient" would want to know), both statements converge on the same question of materiality. Yet, there is currently no consensus on whether use of an AI algorithm is "material information". In the United States, where similar informed consent principles apply, the literature generally suggests that using an AI algorithm to inform healthcare does not require consent. An analogy is sometimes made to physician reliance on professional guidelines and

⁷Richard K Reznick et al, *Task Force Report on Artificial Intelligence and Emerging Digital Technologies*, (Ottawa: Royal College of Physician and Surgeons of Canada, 2020).

⁸ *Loi modernisant des dispositions législatives en matière de protection des renseignements personnels*, LQ 2021, c 25, s 65.2 (not yet in force).

⁹ Office of the Privacy Commissioner of Canada, *A Regulatory Framework for AI: Recommendations for PIPEDA Reform* (Ottawa: 2020), online:

<https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011/>.

protocols, which does not require disclosure. However, Ferron Parayre suggested this comparison might be too simplistic, as protocols and guidelines do not lead to the same kind of personalized decisions as do AI-based predictions. Moreover, she suggested the analogy fails to consider the extent to which physicians may defer to AI, suspending their professional judgement, or the possibility that physician reliance on AI without prior informed consent might undermine patient trust.¹⁰

As a contrasting example, Ferron Parayre discussed recent amendments to France's *Public Health Code* (*Code de la santé publique*) which squarely places on healthcare professionals a duty to disclose the use of AI algorithms to patients (among other obligations).¹¹ She discussed this as an interesting example of a government proactively requiring informed consent, rather than waiting to see whether ordinary informed consent standards will be interpreted to require the disclosure of AI-informed decision-making.

Finally, Ferron Parayre briefly discussed the capacity and informed consent principles applicable to minors, given that this technology is being developed and employed in a children's hospital. She noted that the rules for youth decision-making are different across provinces. In Quebec there is a clear age-based threshold for legal capacity – those 14 years-old or older are presumed to have legal capacity while younger persons require a substitute decision-maker.¹² Other provinces apply different principles. For instance, Ontario law does not include any age threshold for healthcare decision-making rights; it requires a case-specific inquiry into the person's decision-making capacity, which can lead to some persons under 14 years of age having rights to health care decision-making.¹³

¹⁰ On the issue of physicians deferring to AI, Ferron Parayre noted the example of the Therac-25 accidents, in which errors in a software-controlled machine led to six people receiving massive overdoses of radiation. Those operating the machine reportedly noted the unusual volume of radiation but deferred to the machine. (See e.g. NG Leveson & CS Turner, "An Investigation of the Therac-25 Accidents" (1993) 26:7 *Computer* 18; NG Leveson, "The Therac-25: 30 Years Later" (2017) 50:11 *Computer* 8-11.) On the issue of physicians' duties to maintain trust, Ferron Parayre referred to Quebec's *Code of Ethics of Physicians*, CQLR c M-9, r 17 at art 18 ("A physician must seek to establish and maintain with his patient a relationship of mutual trust and refrain from practising his profession in an impersonal manner").

¹¹ Art L4001-3 *Code de la santé publique*.

¹² Art 14 CCQ.

¹³ *Health Care Consent Act*, SO 1996, c 2, Schedule A, s 4. For a discussion of the "mature minor" principle see e.g. *AC v Manitoba (Director of Child and Family Services)*, 2009 SCC 30 at paras 46-69 per Abella J.

B. Privacy (Pierre-Luc Déziel, Université de Laval, Faculté de droit)

Pierre-Luc Déziel spoke to privacy law. He began by noting that the cardiac arrest prediction technology does not present obvious privacy concerns; it doesn't require the collection of new personal information, collected information is stored and used locally, and it is used only for specific purposes – relating to healthcare delivery – that are consistent with the purposes of the original collection. However, he considered some issues that might arise in relation to privacy, noting three such issues: (1) The possible re-identification of patients through membership inference attacks; (2) Bias and the accuracy principle; and (3) Data quantity and the data minimization principle.

With respect to reidentification, Déziel discussed “membership inference attacks”, in which attackers use the algorithm and its internal architecture to reconstruct the previously anonymized training data, gaining access to personal information. There is an obvious privacy issue when this is done to reveal sensitive information such as health data. Déziel considered whether this risk of reidentification should prevent the sharing of AI models, noting that the answer turns, in part, on whether we consider the health information that is imbedded in the model at the training stage to be “personal information”. He explained that guidance on this issue can be found in Ontario’s *Personal Health Information Protection Act*.¹⁴ Section 4 of that Act defines “personal health information” in terms of “identifying information” which is, in turn, defined to include information “for which it is reasonably foreseeable in the circumstances that it could be utilized, either alone or with other information, to identify an individual”.¹⁵ Applying these principles, Déziel suggested the data used to train the cardiac arrest prediction model likely should not qualify as personal health information, given that reidentification is reportedly difficult to achieve. Moreover, reidentification (actual or even attempted) is increasingly seen as an offence under the law, providing a further safeguard. For instance, the privacy law reforms contained in both Quebec’s Bill 64 and the federal Bill C-27, make the identification of persons through deidentified data an offence.¹⁶ Considering both the difficulty of conducting membership inference attacks, and the

¹⁴ *Personal Health Information Protection Act*, 2004, SO 2004, c 3, Sched A.

¹⁵ *Ibid*, s 4.

¹⁶ *An Act to modernize legislative provisions as regards the protection of personal information*, SQ 2021, c 25, s 159; *Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts*, 1st Sess, 44th Parl, 2022 (introduction and first reading, 16 June 2022), online: <<https://www.parl.ca/legisinfo/en/bill/44-1/c-27>> [Bill C-27]. Déziel clarified that Bill C-27 pertains to the private sector and would not apply to this technology, but raised it as an example of how the risk of re-identification can be dealt with, in part, through penal provisions.

possibility of using penal provisions as a dissuasion strategy, Déziel opined that the risk of reidentification should not prevent the sharing of the cardiac arrest prediction model.

A second issue discussed by Déziel was that of selection bias – where an algorithm is trained on data that is not representative of the population on which it will be used. He noted that such bias arguably raises issues not only in discrimination law, but also in privacy law. For guidance, Déziel referred to the Supreme Court of Canada decision in *Ewert v. Canada*.¹⁷ In that case the Correctional Service of Canada had used actuarial risk assessment tools to determine an Indigenous offender’s psychopathy and risk of recidivism. However, the risk assessment tools had been developed with data from non-Indigenous populations. The Court held that this use violated the obligation in the *Corrections and Conditional Release Act* to “take all reasonable steps to ensure that any information about an offender that it uses is as accurate, up-to-date and complete as possible”.¹⁸ Déziel highlighted the interesting nature of this finding – that selection bias can violate an information accuracy principle – given that all Canadian privacy laws contain a similar accuracy principle. For instance, Ontario’s *Personal Health Information Protection Act* requires that health information custodians who are using personal health information “take reasonable steps to ensure that the information is as accurate, complete and up-to-date as is necessary” for the purpose for which it is used, and that they “take reasonable steps to ensure” such accuracy.¹⁹ Reasoning by analogy from the *Ewert* decision, Déziel suggested an AI algorithm that produces a selection bias may contravene not just discrimination law, but also privacy law accuracy requirements. In this context there is arguably apt concern about algorithmic bias among policy makers. For instance, Canada’s Bill C-27 requires that those responsible for “high impact systems . . . establish measures to identify, assess and mitigate the risks of harm or biased output that could result” from those AI systems.²⁰

A last issue discussed by Déziel was that of data quantity and the data minimization principle. He noted a tension between the reality that high quality AI models generally require large amounts of training data, on the one hand, and the well-established privacy law principle that limits collection and use of personal health information to what is “necessary”, on the other. He explained

¹⁷ *Ewert v Canada*, 2018 SCC 30.

¹⁸ *Corrections and Conditional Release Act*, SC 1992, c 20, s 24(1).

¹⁹ *Personal Health Information Act*, 2004, SO 2004, c 3.

²⁰ Bill C-27, *supra* note 16, ss 5, 8. (Déziel noted that the details of these requirements are not yet clear and will need to be clarified in regulation. However, he suggested they signal an awareness that data accuracy is important.)

moreover that “necessity” is interpreted strictly in privacy law; it is not what is useful or practical, but what is essential for achieving the purpose of the collection, use, or disclosure. A key challenge is how to reconcile this minimization principle with the “big data philosophy” that the more data we collect, the better AI will perform.

C. Liability (Vanessa Gruben, University of Ottawa, Faculty of Law)

The third and final legal presentation was given by Vanessa Gruben, who discussed the potential liability issues arising from the cardiac arrest prediction technology. Gruben began by outlining general liability principles, focusing on medical liability in negligence as it applies to health professionals, hospitals, and developers and manufacturers. She explained that any plaintiff seeking to recover damages for medical negligence must establish four elements: (i) the existence of a duty of care between the plaintiff and the defendant; (ii) a violation of the standard of care; (iii) damages; and (iv) a causal relationship between the standard of care violation and the damages. Gruben noted that each of these raises myriad issues but focused her comments on the second element – the standard of care.

Gruben began by discussing the standard of care for health professionals given that they will usually be the ones employing this technology. She explained that liability potentially arises where a health professional makes a “wrong” call, either following an inaccurate assessment from the tool, or choosing to disregard an accurate assessment, resulting in a poor outcome for the patient. Liability might also flow, Gruben noted, where the tool gives a “false positive” result – for instance, wrongly indicating a high risk of cardiac arrest, resulting in clinical interventions that cause patient harm. Gruben explained that to avoid liability, the health professional must show they exercised the level of skill and judgement we would expect from a reasonable health provider in the same context. For instance, if a professional chose to ignore the many other physiological indicators and measurements to instead rely solely on the cardiac arrest prediction tool, and that reliance caused medical harm, the professional might be liable. On the other hand, Gruben noted the expansion of the tool’s purpose – from cardiac arrest prediction to detecting deterioration – which could inject some ambiguity into the standard of care. She suggested it is important to know with precision what physiological indicators are being measured, and for what purpose, to know how this will affect the standard of care. Finally, Gruben explained that if use of this tool eventually becomes standard practice, this could become part of the standard of care such that failure to use it, in a situation that leads to patient harm, could attract health professional liability.

Gruben also discussed the potential for hospital liability where use of the predictive tool leads to patient harm. She explained that this liability may be direct; hospitals owe a range of duties to patients, for instance to implement safe systems, and provide adequate training. In this context, a hospital could be liable if it failed to appropriately vet, or monitor use of, the predictive tool, which led to patient harm. Hospitals can also be vicariously liable for the negligence of their employees including nurses and other staff when using the predictive tool. (An exception, Gruben explained, relates to physicians who are usually considered to be independent contractors rather than employees, such that hospitals are not vicariously liable for their negligent acts).

Finally, Gruben explained that developers and manufacturers can be liable for design flaws in their products. In common law jurisdictions, patients can bring negligence claims directly against product developers. The standard here is whether the developer used “reasonable care” in the circumstances; if they relied on flawed data or there was an error in the model, the developer could be liable for resulting harm. Gruben also noted that developers and manufacturers can be liable if they know of a risk and fail to warn users of the product of that risk. This is, moreover, an ongoing duty; if new concerns arise, it is incumbent on the developer to proactively disclose them. There are some defences to product liability claims – for instance, relating to misuse of the product – however this remains an area of potential liability risk for developer / manufacturers. On the other hand, Gruben noted the possibility that developers and manufacturers will seek to limit their liability through contract (e.g., liability exclusion clauses), potentially transferring more risk to hospitals and healthcare providers.

Breakout Sessions



The commentaries were followed by breakout sessions on each of privacy, informed consent, and liability. The purpose of these was to allow members to more deeply engage with the potential legal issues – and possible solutions – raised by the technology. A rapporteur from each group summarized the discussion and findings and presented these in a debriefing session with the full group.

The breakout sessions were intended to – and did – generate the kind of creative thoughts and insights made possible by deep engagement by an interdisciplinary group. While this report cannot include all the points discussed, some core thematic concerns that arose in each session are summarized below.

Breakout #1: Informed Consent

Attendees: Catherine Régis (Rapporteur), Caroline Mercer (Scribe), Philip Mathew, Sylvain Bédard, Sana Tonekaboni, Audrey Ferron Parayre, Ian Stedman

Members of this breakout group started by asking what kind of information must be disclosed, as part of informed consent, and why. Many were in general agreement that this cardiac arrest prediction technology is currently low-risk, such that there might be a low or even *no* requirement for informed consent prior to its use. At the same time, they discussed the possibility that automation bias could alter this situation – that is, if clinicians begin to defer to the tool, without exercising independent judgement, use of the technology could require informed consent. (They gave the example of Google maps; people know they have the choice to ignore the program's directions but may follow them without exercising any judgement.)

The group also discussed several important gaps between current law and practice and what might be optimal. One such gap is between the *law* of informed consent and how those principles are carried out in *practice*. They noted that despite strict legal requirements, many medical procedures already take place without meaningful consent. They discussed the need to confront the “on paper” nature of informed consent to make it more genuine. Another potential gap mentioned was the one between current law and a more *ideal* law of informed consent. A member

noted the connection of informed consent law to ideas of negligence; the law requires disclosing a proposed medical treatment's material "risks and benefits", which are also key to thinking about liability for harm. The member suggested a better approach might also consider what information is necessary to ensure transparency and build trust. (She noted that to some extent this principle applies already; if a patient is asking questions, the physician must provide meaningful answers as part of the informed consent process, whether or not they relate to risks and benefits.)

The group also considered whether informed consent could or should require disclosing the extent to which the tool is appropriate for that patient – for instance, whether it was trained on data with representative ethnicity. They acknowledged this kind of demographic information is not currently collected by the healthcare system, raising questions about how healthcare professionals can know the limits of the tool. Finally, members of this group considered whether informed consent principles would or should be applied uniformly across patients or whether perhaps a subset of patients (e.g., those presenting with higher risk) should be required to consent. One member suggested consent might be required where the tool flags a high risk and clinicians plan to intervene. Another participant noted, on the other hand, that the technology can lead to false negatives, causing clinicians to focus elsewhere, which can be equally risky.

Breakout #2: Privacy

Attendees: Michael Da Silva (Rapporteur), Nicole Davidson (Scribe), Bryan Thomas, Christopher Viney, Lindsay Thompson, Natasha Ovtcharenko, Pierre-Luc Deziel, Melissa McCradden

Members of the privacy breakout group discussed three main issues: (i) the fact that not all AI raises unique moral and legal issues; (ii) the risk of reidentification; and (iii) possible data sharing problems. On the first issue, they suggested this tool is a reminder that some AI technologies are relatively unproblematic from a legal and ethical point of view. They considered this prediction tool to be one such technology, which is very low risk from a privacy standpoint, given that it operates within a single institution that already has rigorous privacy protections in place.

Regarding reidentification, the group queried whether this is a genuine risk, with some members expressing skepticism. Indeed, some members wondered whether the criminalization of data reidentification (which is being pursued in some jurisdictions) addresses a real problem or, conversely, might create unwarranted fear about data security and undermine trust in the long term. The group also considered the possible privacy risks arising from sharing the model and associated data. Despite having considered the risk of reidentification to be low, the group discussed the possibility that some third parties – such as insurance companies – might be especially motivated to access sensitive data. On the other hand, group members suggested that some data sharing could be positive, for instance, where other developers and manufacturers want to develop complementary healthcare technologies. This group also briefly discussed other privacy issues, including data ownership (whether people should own their own data); whether privacy impact assessments should be required even for low-risk tools; the possible trade-offs between privacy and having high-quality data for AI development; and data representativeness. On the last issue, the group was not very concerned about possible selection bias in the model's training data; members suggested that concerns that unrepresentative data might undermine the model's performance could be investigated and resolved during testing. (Though a workshop member challenged this, noting there are issues with asking people to submit, in real time, to a tool that was not trained with data that includes their demographic.)

Breakout #3: Liability

Attendees: Amy Zarzeczny (Rapporteur), Saly Sadek (Scribe), Michael Froomkin, Genevieve Lavertu, Colleen Flood, Jennifer Kingdon, Cécile Bensimon, Vanessa Gruben

The breakout group on liability discussed several important issues. One was the intersection of responsibility and liability for appropriate AI use with the reality of uneven resources. Members questioned, for instance, what should be done if this predictive tool becomes part of the standard of care but is inaccessible to some regions or institutions. They acknowledged that resource constraints are a broader issue but suggested this should be a particular focus in thinking about AI development and deployment. A related issue raised was that of safe implementation and upkeep of AI technology. Members discussed the possibility that a jurisdiction or institution might adopt an AI technology but lack the resources or expertise to make the continuous calibrations and adjustments that are necessary to ensure ongoing safety and effectiveness. The group discussed the particular importance of specifying institutional oversight responsibilities, especially if AI tools like this one are considered to fall outside the regulatory ambit of Health Canada.

Another issue discussed by this group was the effect of an evolving product on liability. They highlighted the predictive tool's shifting purpose – from cardiac arrest prediction to predicting deterioration – along with the possibility of further evolutions. They noted that such shifts can create uncertainty about the standard of care and related risk of liability. They also noted the potential for developers to try to limit their own liability through contractual limitation clauses and waivers. The group suggested this also is an area that might require regulatory oversight.

Members of this group also discussed issues relating to explainability – that is, the challenges that potentially arise when AI algorithms conduct “black-box” reasoning that cannot be scrutinized. They queried what effect this might have on liability, hypothesizing that some degree of explainability might be important to a physician who exercises judgement to depart from the tool's assessment or recommendation. A lack of AI explainability might make it more difficult for a physician to justify such a decision, increasing their risk of liability in cases of patient harm.

As a general matter, the group discussed the potential need for regulatory oversight of AI in healthcare. They highlighted the need for review of current regulatory structures, including the role of Health Canada, with a view to possible reform. They suggested, moreover, that such

discussions should take into account the evolving perceptions, for instance, regarding AI's purpose, effectiveness, and safety, held by those affected (including providers, patients, and the public). At the same time, the group emphasized the need for nuance, noting the heterogeneity of AI technologies and their legal and ethical implications.

DISCUSSION

I. Interplay between the different issues

One important theme was the tensions and interrelationships between the different legal issues. For instance, the privacy protection that is provided by data minimization, among other privacy principles, may reduce the ability of developers to create algorithms that are safe and effective (raising possible liability issues) and representative (raising potential discrimination and, paradoxically, privacy issues, given the possibility that selection bias contravenes privacy law's accuracy principle). A biased data set could also implicate informed consent; if a tool is trained on data that is not representative of a patient's demographic, with an attendant risk of harm, there might be a duty to disclose that fact and obtain the patient's informed consent. This kind of use might also violate the health professional's and/or the institution's duty of care to ensure the adequacy of the tools used in patient care. Déziel noted that the issue in the *Ewert* case was perhaps not just that the Correctional Services of Canada used an actuarial tool that was trained on a different population, but that it did so knowingly, without any attempt to adapt it to the individual.

On the other hand, *failure* to appropriately protect privacy (and perhaps also address data ownership questions) can also undermine the autonomy and patient trust that underlies principles of informed consent. It can also give rise to real harms for which health professionals, health institutions, and developer/manufacturers may be liable. The intention here is not to single out privacy; similar connections and interrelationships could be identified beginning from any of the relevant legal issues. The point is that many of these issues are in tension; resolving their appropriate balance and treatment likely requires nuanced analysis and, in some cases, regulatory oversight.

II. The challenge of informed consent

Another theme that warrants emphasis is the challenging nature of informed consent.

Participants were divided on what should be the legal standard and how to ensure meaningful operation of this standard in practice, questions that apply beyond the AI context. While each of the legal issues received vigorous treatment, this issue was among the most challenging for participants.

Participants differed on how to characterize the cardiac prediction technology. Some emphasized its similarity to a cardiac or oxygen monitor, the use of which doesn't require specific informed consent. They noted that when parents admit their children to the ICU they understand and expect that their children will be on monitors; such equipment isn't usually consented to unless it will be implanted in the patient's body. If the cardiac prediction tool is like a heart monitor, and simply offers existing data flows in a new format, it can arguably be used without consent. Some bolstered this position by emphasizing the "human in the loop" – that is, the role of health professionals in interpreting the tool's signals and deciding what treatment should follow. Describing the tool in this way, one participant stated simply, he was "not feeling the nervousness" around informed consent.

Others were less ready to wholly accept the "cardiac monitor" analogy and argued for a more cautious approach to informed consent. One of the core concerns, for these participants, was that clinicians would over rely on the tool, thus failing to provide human oversight and judgement. One reason for this concern was automation bias; as one participant put it, "what if physicians feel they *must* follow the machine?" Another was the reality of limited resources in stressful ICU environments. Where a machine provides a risk analysis that apparently relieves some of the burden on overstretched health professionals they will, one participant suggested, "almost certainly rely on the machine". Indeed, this person viewed the notion of physicians in busy ICU environments taking in all the information, interpreting the AI-generated risk assessment, knowing the machine's limitations, and then providing their own judgement, as somewhat mythical. Others disagreed and argued clinician overreliance should be considered a violation of the standard of care, triggering negligence. On the other hand, there was general agreement that if future evolutions of the technology allow it to act autonomously *by design*, this use would require informed consent.

Another aspect of this disagreement was seemingly the appropriate level of caution. A participant emphasized that this technology – like *all* AI-based tools – is new. Whatever analogy (to another monitor, to another member of the team) might currently seem appropriate, we can't know with

certainty how the tool will be used at the point of care and as it evolves. Moreover, this participant noted, there are historical examples of AI healthcare tools being designed to consider interests beyond those of the patient. She gave the example of a technology that guides the timing of patient discharge from hospital that is programmed to also consider cost effectiveness. Given potential competing interests, among other factors, this participant preferred France's approach of erring on the side of telling patients AI is being employed.

III. The law - practice gap and the importance of patient perspectives

A final overarching point that warrants emphasis is the potential law-practice gap and the need to attend to patient perspectives. This point overlaps with – but also goes beyond – the discussion of informed consent.

As stated previously, there was significant discussion of the challenge of obtaining meaningful informed consent in practice. Participants noted that even now, patients needing care are presented with long and complicated consent forms, which they will often sign without true understanding. Where AI brings added complexity, some participants were skeptical about the possibility of meaningful informed consent involving AI. These discussions suggest a possible need for law and practice reform; however, they also highlight the importance of considering the humans at the heart of this enterprise. For instance, participants noted that the experience and outcomes of informed consent processes will vary significantly with how the clinician of the day characterizes the information, and with the patient's situation and values (e.g., their own perceptions of the urgency of care). A participant also wondered whether AI could sometimes be *empowering* – moving more (and more meaningful) information into the hands of patients, which they could then discuss with their physicians. They noted some physicians might not like this idea (of patients coming with their own and perhaps even competing information) but this kind of patient perspective arguably requires consideration.

Drawing from another area of law – privacy – participants again asked pertinent questions about patient perspectives. For instance, they asked whether the criminalization of data de-identification might increase willingness to share data, on the one hand, or lead to unwarranted fear of hacking and inference attacks, on the other. A participant also suggested that this kind of perspective – what patients want and need – could usefully drive more AI development. She noted that current development usually begins from needs that are identified by healthcare professionals. However,

this process could be more patient-driven, for instance, by having patient coalitions identify problems with the healthcare system as experienced by patients.

Conclusion



This workshop complemented and built on previous ones, relating to the “OR Black Box”, the “Suicide Artificial Intelligence Prediction Heuristic”, and “digital twins” technology.²¹ It reinforced once more that looking at particular use-cases of AI in healthcare, and moreover doing so with an expert multidisciplinary group, provides needed nuance and particularity; it allows for deep exploration of the particular legal and ethical issues that arise from a given technology, recognizing the heterogeneity of both AI technologies and their associated challenges. The analysis of the cardiac arrest prediction technology brought to light a number of such issues – relating to privacy, informed consent, and liability. It also allowed for an in-depth conversation about the ability of current law to address these challenges along with the possible need for law reform. At the same time, this *micro* perspective is complemented by a *macro* one; examining the individual technologies through a series of case studies brings to light cross-linking issues, problems, and possible solutions. This dual perspective can provide essential guidance for both AI developers and lawmakers seeking to determine the optimal directions for regulatory reform.

²¹ See *supra* note 2.

Workshop Participants:

Adina Panchea	François Ferland	Pascal Thibeault
Amy Zarzeczny	Gagan Gill	Philip Mathew
Anna Goldenberg	Genevieve Lavertu	Pierre-Luc Déziel
Bryan Thomas	Hazar Haidar	Regiane Garcia
Audrey Ferron Parayre	Ian Stedman	Saly Sadek
Caroline Mercer	Jennifer Kingdon	Sana Tonekaboni
Catherine Régis	Joanne Kim	Shanil Keshwani
Catherine Régis	Lindsay Thompson	Sophie Nunnelley
Cécile Bensimon	Michael Da Silva	Sylvain Bedard
Christopher Viney	Melissa McCradden	Tanya Horsley
Colleen Flood	Natasha Ovtcharenko	Vanessa Gruben
Elissa Strome	Nicole Davidson	Wendy Halle (CIFAR Events Team)
Emily Nicholas Angl		



CIFAR

cifar.ca/ai