Governing the Real World Application of Medical AI

Conference Report

15 - 16 Nov 2024

Hong Kong & Zoom

Disclaimer:

All the statements and information/ suggestions in this publication are made/ provided without legal responsibility and do not constitute legal or other professional advice. No legal liability shall arise from any errors or omissions in relation to any of the statements/ information/ suggestions in this publication. No express or implied warranties of accuracy or fitness for any particular purpose or use with respect to the statements/ information/ suggestions in this publication are made. The opinions expressed by the contributors in this publication do not necessarily reflect the views of the University of Hong Kong. This publication is for general reference only. The law changes from time to time.

Organiser:

CMEL
THE UNIVERSITY OF HONG KONG CENTRE FOR MEDICAL ETHICS AND LAW

Supporting Organisations:

HONG KONG ACADEMY OF MEDICINE
香港醫學專科學院

HKU Med School of Clinical Medicine
Medical Ethics & Humanities Unit
香港大學醫學倫理及人文學部

About This Report

The Centre for Medical Ethics and Law (CMEL) of the University of Hong Kong (HKU) develops new ideas and solutions in response to the big ethical, legal and policy questions of medicine and health. CMEL was founded in 2012 by the LKS Faculty of Medicine and Faculty of Law at HKU. One of CMEL's flagship research areas is the regulation of health technologies.

Artificial Intelligence (AI) has the potential to transform and revolutionise healthcare. Numerous AI applications in healthcare have been demonstrated to be highly accurate and conducive to a range of clinical applications in terms of safety and diagnosis.[1]

The use of AI in healthcare is expected to bring significant benefits, but it may also give rise to uncertainties and risks. The nature and magnitude of the risks associated with such use of AI might not be fully understood at the moment. Proper governance of AI in healthcare is of paramount importance. It can facilitate innovation, protect patients, and support healthcare professionals.

In mid-2024, CMEL saw the need to participate in the dialogue to examine which approach to governing AI in healthcare would be the most optimal for the Hong Kong Special Administrative Region (HKSAR or Hong Kong) of the People's Republic of China and to build a platform for this purpose. Consequently, CMEL hosted a conference titled "Governing the Real World Application of Medical AI" in Hong Kong on 15 and 16 November 2024 to explore this issue from legal and ethical perspectives and to facilitate sharing and learning from others' experiences.

At the conference, the presenters and delegates discussed the governance of AI in healthcare in Hong Kong, the Macau Special Administrative Region (Macau SAR or Macau), Mainland China, Europe (the European Union, the Council of Europe, and the United Kingdom), Canada, the United States, and the Middle East, showcasing a diverse range of governance approaches. Non-jurisdiction-specific topics such as the principles of governance and ethical challenges were also discussed. The conference programme and the biographies of the contributors are available in the appendices.

This report presents expanded abstracts submitted by speakers after the conference and, in certain cases, transcripts/summaries of their presentations. These materials have been re-organised to better present different approaches to the governance of AI in healthcare around the world vis-à-vis the situation in Hong Kong. The report concludes with an over-arching "Analysis" of the above materials as well as the potential way forward for the governance of healthcare AI in Hong Kong.

CMEL hopes that this report will help policymakers and healthcare providers chart the way forward in this new landscape of great opportunities and uncertainties. As AI technology develops, frequent review and refinement of the governance approach are necessary to ensure its effectiveness.

Prof Gilberto Leung
Co-Director, Centre for Medical Ethics and Law
Tsang Wing-Hing Professor in Clinical Neuroscience, Clinical Professor and Director of the School of Clinical Medicine of the LKS Faculty of Medicine
The University of Hong Kong

[1] Levine, David M., Rudraksh Tuwani, Benjamin Kompa, *et al*., "The Diagnostic and Triage Accuracy of the GPT-3 Artificial Intelligence Model: An Observational Study", *The Lancet Digital Health* 6, no. 8 (2024): e555–e561, e555. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(24)00097-9.

Prof Eric C. Ip
Co-Director of the Centre for Medical Ethics and Law
Professor of the Faculty of Law
The University of Hong Kong

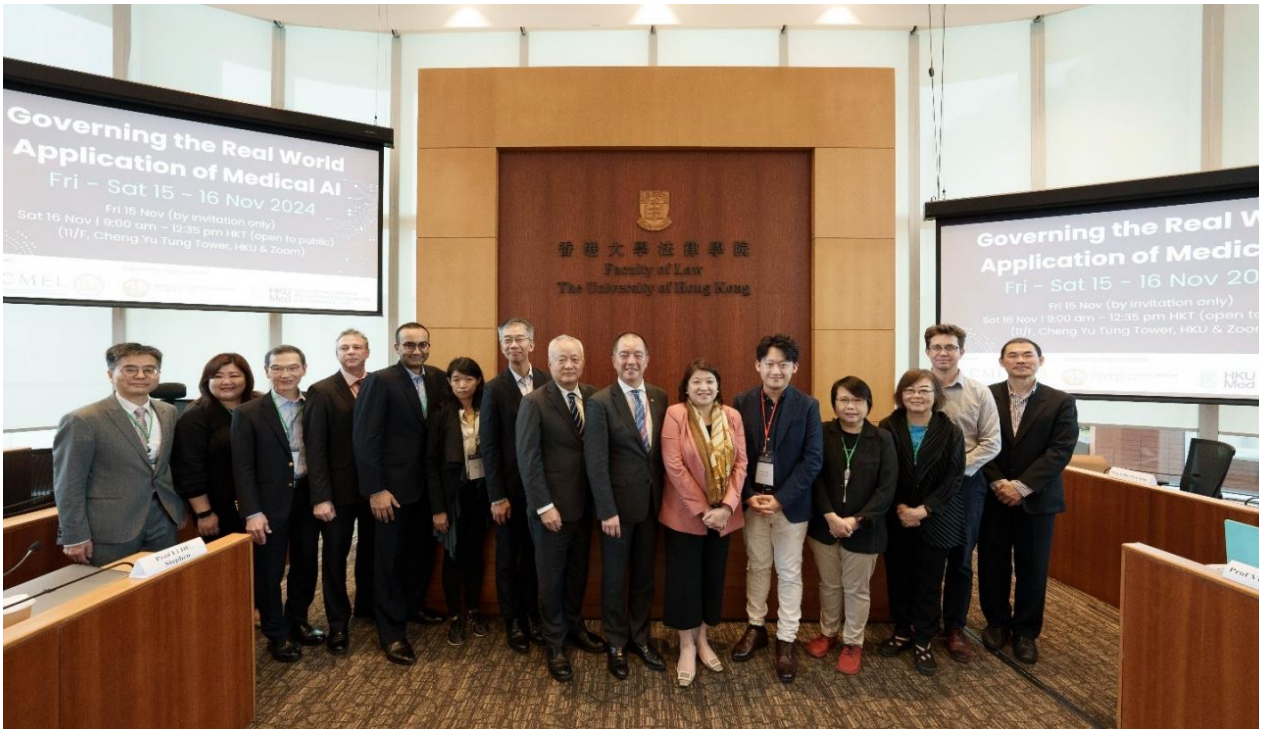
Prof Julie Chen
Director, Medical Ethics and Humanities Unit (supporting organisation of the conference), School of Clinical Medicine
Associate Professor of Teaching and Assistant Dean (Student Wellness & Engagement) of the LKS Faculty of Medicine
The University of Hong Kong

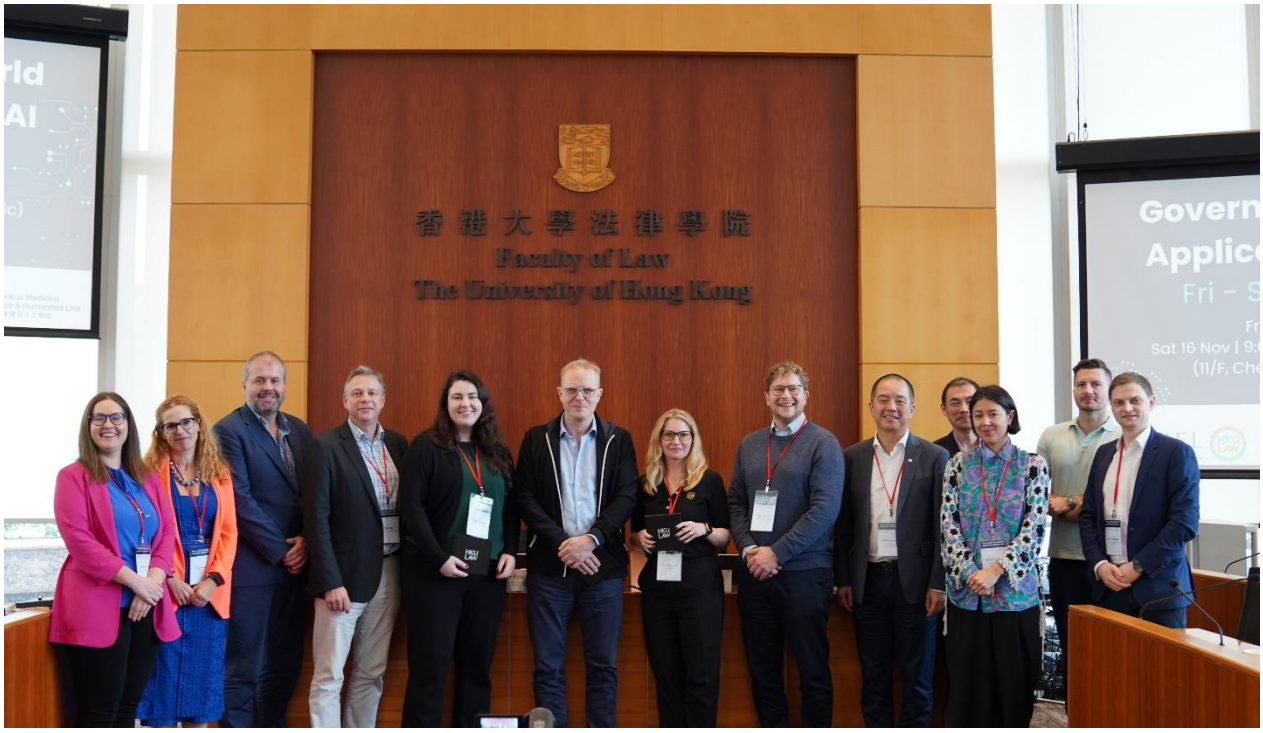
Dr Neeraj Mahboobani
Convenor of the Task Force on Artificial Intelligence of the Professionalism and Ethics Committee of the Hong Kong Academy of Medicine (supporting organisation of the conference)

Dr Pete Mills
Director, PHG Foundation (collaborating institution for the conference), University of Cambridge, United Kingdom

May 2025

Photos taken at the conference





Executive Summary

It is widely believed that Artificial Intelligence (AI) will revolutionize healthcare. However, the rapid evolution of these technologies has also given rise to challenges and uncertainties in regard to the governance of AI in healthcare and other fields.

This report contains expanded abstracts on the governance of AI in healthcare in the Hong Kong Special Administrative Region, the Macau Special Administrative Region, Mainland China, Europe (the European Union, the Council of Europe, and the United Kingdom), and the Middle East as well as a presentation transcript and a presentation summary on the governance of AI in healthcare in the United States. Although this report does not contain any expanded abstract, presentation transcript or presentation summary on the governance of AI in healthcare in Canada, the same is discussed in the "Analysis" section of this report.

The European Union (EU) has enacted an AI Act to comprehensively regulate AI across sectors. The EU AI Act, *inter alia*, prohibits certain AI practices and imposes regulatory requirements on AI systems and general-purpose AI models. The EU Act expressly excludes certain AI systems, certain AI models and certain entities from its scope of application. For example, Article 2 of the EU AI Act provides that this Regulation does not apply to, among others, AI systems or AI models specifically developed and put into service for the sole purpose of scientific research and development. It should be noted that the EU AI Act has an extraterritorial effect in some circumstances.

In contrast to the approach taken by the EU, some of the other jurisdictions included in this report have adopted a pro-innovation approach to the governance of AI (e.g., United Kingdom, United States), while the other jurisdictions have adopted a middle-ground approach (e.g., Canada).

Despite the diversity of approaches adopted by these jurisdictions, there was broad agreement around certain values and considerations that ought to underpin the governance of AI, such as the importance of adopting an adaptive approach that safeguards against harm, aligning AI governance with medical device regulation, governing AI throughout the life cycle, regulating products and systems, global coordination and regulation embedding ethical values.

In regard to the situation in Hong Kong, several guidance documents on AI have been issued by relevant public authorities. The Digital Policy Office (DPO) of the Government of the HKSAR issued the "Ethical Artificial Intelligence Framework" and the "Hong Kong Generative Artificial Intelligence Technical and Application Guideline" in 2021 and 2025, respectively. The Office of the Privacy Commissioner for Personal Data (PCPD) of Hong Kong published the "Guidance on the Ethical Development and Use of Artificial Intelligence", "Artificial Intelligence: Model Personal Data Protection Framework" and "Checklist on Guidelines for the Use of Generative AI by Employees" in 2021, 2024, and 2025, respectively. As for the regulatory landscape for AI medical devices, please refer to the expanded abstract of Engineer Lam Kam Chun, Tommy, of the Department of Health of the Government of the HKSAR in this report. Lastly, it should be noted that the Intellectual Property Department (IPD) of the Government of the HKSAR conducted a public consultation in 2024 to seek public views on deepfakes, the transparency of AI systems, and certain copyright issues relating to AI.

Acknowledgements

We are extremely grateful to the non-local speakers for travelling to Hong Kong for the conference and to the local organisations for sending representatives to attend the conference as speakers and/or delegates.

We would like to express sincere gratitude to WYNG Foundation for their funding support to CMEL and to the following representatives from WYNG Foundation in particular:

- Ms Anna Hung-yuk Wu, Trustee of WYNG Foundation (as she then was at the time of the conference), CMEL Executive Board Member (as she then was at the time of the conference) and Honorary Professor of the HKU Faculty of Law
- Ms Glenda Yu, CEO of WYNG Foundation

We would like to thank the following organisations and individuals for their participation at the conference:

- Health Bureau of the Government of the Hong Kong Special Administrative Region (HKSAR), in particular:
  - Dr LEE Ha Yun, Libby, JP, the Under Secretary for Health of the Health Bureau
  - Dr TO May Kei, Liza, Assistant Director (Health Sciences and Technology) of the Department of Health
  - Dr CHAN Kwok Hung, Addi, Principal Medical and Health Officer (Medical Device) of the Department of Health
  - Engineer LAM Kam Chun, Tommy, Senior Electronics Engineer (Medical Device) of the Medical Device Division of the Department of Health
- Digital Policy Office of the Innovation, Technology and Industry Bureau of the Government of the HKSAR, in particular:
  - Mr Donald Mak, Deputy Commissioner (Data Governance) of the Digital Policy Office
- Hong Kong Academy of Medicine ("HKAM"), one of the supporting organisations of the conference, in particular:
  - Dr James Shing-ping Chiu, Co-Chairman of the Professionalism and Ethics Committee of HKAM
  - Dr Neeraj Mahboobani, Convenor of the Task Force on Artificial Intelligence of the Professionalism and Ethics Committee of HKAM
- HKU Medical Ethics and Humanities Unit ("MEHU"), one of the supporting organisations of the conference, in particular:
  - Prof Julie Chen, Director of MEHU of School of Clinical Medicine, Associate Professor of Teaching and Assistant Dean (Student Wellness & Engagement) of HKU LKS Faculty of Medicine
- HKU Faculty of Law, in particular:
  - Dean, Prof Hualing Fu, and the support staff
- HKU LKS Faculty of Medicine, in particular:
  - Dean, Prof Chak-sing Lau and the support staff
- Prof Calvin Wai-Loon Ho, Associate Professor of Faculty of Law at Monash University, CMEL Research Fellow and immediate past CMEL Co-Director
- PHG Foundation of the University of Cambridge, in particular:
  - Dr Pete Mills, Director of the PHG Foundation
  - Dr Colin Mitchell, Head of Humanities of the PHG Foundation
  - Ms Tanya Brigden, Senior Policy Analyst (Biomedical Ethics) of the PHG Foundation
  - Mr Jamie Hearing, Policy Analyst (Humanities) of the PHG Foundation

- Hong Kong Genome Institute, in particular:
  - Dr LO Su-vui, CEO of the Hong Kong Genome Institute
- Medical Protection Society ("MPS"), in particular:
  - Dr Danny Lee, Senior Medical Advisor of MPS
  - Ms Joeky Leung, Regional Development Director for Asia of MPS
- all of the conference participants and helpers.

Special thanks must go to Dr Pete Mills, Dr Colin Mitchell, Ms Tanya Brigden and Mr Jamie Hearing of the PHG Foundation for assisting us in organising the conference and, most notably, in drafting the section of "Analysis" in this report. We are immensely grateful for their insights and invaluable contributions.

Table of Content

Part 5: US

Part 6: Middle East

Part 7: Not Jurisdiction-specific

Disclaimer:

All the statements and information/ suggestions in this publication are made/ provided without legal responsibility and do not constitute legal or other professional advice. No legal liability shall arise from any errors or omissions in relation to any of the statements/ information/ suggestions in this publication. No express or implied warranties of accuracy or fitness for any particular purpose or use with respect to the statements/ information/ suggestions in this publication are made. The opinions expressed by the contributors in this publication do not necessarily reflect the views of the University of Hong Kong. This publication is for general reference only. The law changes from time to time.

Part 1: Hong Kong SAR – 1

Opportunities and Challenges of Applying AI for Medical Applications

Mr Donald Mak
Deputy Commissioner (Data Governance)
Digital Policy Office, Innovation, Technology and Industry Bureau
Government of the HKSAR, China

Introduction

Artificial Intelligence (AI) is becoming a key driver of innovation in the digital era, which can also significantly benefit the medical field. Its rapid advancement presents exciting opportunities to revolutionise healthcare by improving diagnosis, personalising treatment plans, and enhancing patient outcomes. However, these advancements also bring about challenges, which demand for an AI ecosystem that can balance the opportunities and risks brought about by AI. The Government of the Hong Kong Special Administrative Region ("the Government") is adopting a multi-pronged approach to develop an AI ecosystem, with a view to enhancing public services and driving economic growth.

Opportunities of AI in Healthcare

The integration of AI into healthcare is revolutionising medical services. AI systems can analyse vast amounts of medical data, identifying patterns that enhance diagnosis and treatment planning, leading to improved patient outcomes. For instance, AI algorithms can quickly analyse complex medical images, detecting subtle indicators of disease that may be overlooked by humans.

Additionally, AI excels in predictive analytics, forecasting disease outbreaks and patient deterioration, which enables early intervention and better resource allocation. In drug discovery, AI accelerates the identification of promising candidates and predicts their interactions and side effects. Furthermore, AI enhances patient engagement through personalised health recommendations and automates routine tasks like scheduling, allowing healthcare professionals to focus more on patient care.

Challenges and Ethical Considerations

The application of AI in healthcare presents numerous challenges and ethical considerations that must be addressed to ensure effective and responsible use. AI systems require vast amounts of health data for training, often involving sensitive information, which raises significant concerns about privacy and security.

Bias can arise from unrepresentative training data, leading to discriminatory outcomes and resulting in disparities in diagnosis and treatment recommendations. Human oversight is therefore crucial, as AI systems can make errors or exhibit biases, necessitating human intervention to validate and interpret their results. The integration of AI in healthcare also raises complex questions about accountability and liability, particularly when AI-driven decisions lead to adverse results.

Furthermore, some AI systems operate in ways that make it challenging to understand their decision-making processes. This can impact trust among healthcare professionals and patients. Ensuring the reliability and robustness of AI systems is critical for achieving consistent performance across diverse clinical settings, requiring extensive validation and testing.

Government's Pro-innovation Approach in Supporting AI Development

The Government is mindful of the associated challenges and ethical concerns of AI. In response, the Government has adopted a comprehensive strategy to develop the AI ecosystem, which encompasses measures to continuously improve AI governance, enrich data resources, establish the necessary digital infrastructure, foster talent development, promote research and development, and facilitate the adoption of AI.

Ethical Artificial Intelligence Framework[2]

The Government formulated the "Ethical Artificial Intelligence Framework" in 2021, which consists of principles, practices, and assessments designed to assist Government bureaux and departments (B/Ds) in AI adoption, ensuring that ethical considerations are integrated into the planning, design, and implementation of AI projects.

Under the Framework, there are twelve ethical AI principles. Two key principles—(1) Transparency and Interpretability, and (2) Reliability, Robustness, and Security—are classified as "Performance Principles", which are essential for the successful implementation of the other principles.

The remaining principles are categorised as "General Principles," which include (1) Fairness, (2) Diversity and Inclusion, (3) Human Oversight, (4) Lawfulness and Compliance, (5) Data Privacy, (6) Safety, (7) Accountability, (8) Beneficial AI, (9) Cooperation and Openness, and (10) Sustainability and Just Transition.

Notably, the Transparency and Interpretability principle requires organisations to clearly explain AI decision-making processes, which is essential for fostering trust in healthcare sector. The Reliability, Robustness, and Security principle ensures that AI applications function reliably and securely, which is crucial for high-risk applications in the medical field. Lastly, the Human Oversight principle emphasises informed decision-making by users, ensuring that human expertise remains central to patient care and mitigating risks associated with AI errors.

Hong Kong Generative Artificial Intelligence Technical and Application Guideline[3]

The Digital Policy Office (DPO) also commissioned the "Hong Kong Generative AI Research and Development Center" (HKGAI), which was established with the funding support of the AIR@InnoHK platform, to study and propose appropriate codes and guidelines on the generative AI technology. With the completion of this study, the DPO promulgated the "Hong Kong Generative Artificial Intelligence Technical and Application Guideline" ("the Guideline") in April 2025 with a view to establishing a governance framework on the use of generative AI technology that can better fit in the environment of Hong Kong, balancing AI innovation, application and responsibility.

The Guideline aims to provide practical and operational guidance for technology developers, service providers, and service users in the application of generative AI technology. It covers the scope and limitations of applications, potential risks and governance principles of generative AI technology.

---

[2] Digital Policy Office of the Government of the Hong Kong Special Administrative Region of the People's Republic of China ("The Government of the HKSAR"), "Home: Our Work: Data Governance: Enhancing Data Governance: AI and Data Ethics", Ethical Artificial Intelligence Framework (Customised Version for General Reference by Public) Version: 1.4, 25 July 2024, https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/.

[3] Digital Policy Office of the Government of the HKSAR, "Hong Kong Generative Artificial Intelligence Technical and Application Guideline", April 2025, https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/doc/HK_Generative_AI_Technical_and_Application_Guideline_en.pdf.

The Guideline features a four-tiered risk classification, with a proportionate management approach based on potential risk. The applications posing existential threats (such as causing harm and affecting human life) should be classified as "Unacceptable Risk" and should be prohibited, with the developers held liable. "High Risk" applications, such as those in healthcare diagnostics or autonomous driving, should require conformity assessments and human oversight. "Limited Risk" applications, like recruitment tools, educational AI, should meet transparency obligations and undergo annual audits. Finally, for "Low Risk" applications, like email spam filters, creativity tools, self-certification should be sufficient.

Safeguarding Personal Data Privacy

Hong Kong has placed significant emphasis on safeguarding data privacy. In 2021, the Office of the Privacy Commissioner for Personal Data (PCPD), an independent body overseeing the implementation of and compliance with the Personal Data (Privacy) Ordinance (PDPO) in Hong Kong, issued the "Guidance on the Ethical Development and Use of Artificial Intelligence".[4] In June 2024, the PCPD established the "Artificial Intelligence: Model Personal Data Protection Framework", outlining how AI should be procured, implemented, and utilised in accordance with the PDPO.[5] In March 2025, PCPD issued the "Checklist on Guidelines for the Use of Generative AI by Employees"[6] for employees at work complying with the requirements of the PDPO in relation to the handling of personal data.

Sector-specific Regulations

Various regulatory bodies in Hong Kong oversee different industries, implementing specific regulations, guidelines, and codes of conduct. Sector-specific guidelines can help to address the unique challenges of specific sectors, striking a balance between innovation and ethical considerations.
For instance, the Hong Kong Monetary Authority (HKMA) has published the "High-level Principles on Artificial Intelligence"[7], "Consumer Protection in respect of Use of Generative Artificial Intelligence"[8] and "Use of Artificial Intelligence for Monitoring of Suspicious Activities"[9].

[4] Office of the Privacy Commissioner for Personal Data of Hong Kong, "Home: Resources Centre: Publications: Guidance Notes/ Reports", Guidance on the Ethical Development and Use of Artificial Intelligence (August 2021), effective August 2021, https://www.pcpd.org.hk/english/resources_centre/publications/guidance/guidance.html?year=2021.
[5] Office of the Privacy Commissioner for Personal Data of Hong Kong, "Home: Resources Centre: Publications: Guidance Notes/ Reports", Artificial Intelligence: Model Personal Data Protection Framework (June 2024), effective June 2024, https://www.pcpd.org.hk/english/resources_centre/publications/guidance/guidance.html?year=2024.
[6] Office of the Privacy Commissioner for Personal Data of Hong Kong, "Home: Resources Centre: Publications: Guidance Notes/ Reports", Checklist on Guidelines for the Use of Generative AI by Employees (March 2025), effective March 2024, https://www.pcpd.org.hk//english/resources_centre/publications/files/guidelines_ai_employees.pdf
[7] Hong Kong Monetary Authority, "Home: Regulatory Resources: Regulatory Guides: Circulars: Archive: 2019", High-level Principles on Artificial Intelligence, 1 November 2019, https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2019/20191101e1.pdf.
[8] Hong Kong Monetary Authority, "Home: Regulatory Resources: Regulatory Guides: Circulars: Archive: 2024", Consumer Protection in Respect of Use of Generative Artificial Intelligence, 19 August 2024, https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2024/20240819e1.pdf.
[9] Hong Kong Monetary Authority, "Home: Regulatory Resources: Regulatory Guides: Circulars: Archive: 2024", Use of Artificial Intelligence for Monitoring of Suspicious Activities, 9 September 2024, https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2024/20240909e1.pdf.

The Financial Services and the Treasury Bureau (FSTB) issued the "Policy Statement on Responsible Application of Artificial Intelligence in the Financial Market"[10] outlining the Government's policy stance and approach towards the responsible use of AI in the financial sector.

Open Data and Sharing of Data for AI Development

Effective data governance plays a strategic role in driving our economy towards digital transformation, high-quality development and healthy development of AI. The Innovation, Technology and Industry Bureau (ITIB) published the "Policy Statement on Facilitating Data Flow and Safeguarding Data Security in Hong Kong"[11] in December 2023 to set out the Government's management principles and key strategies as well as 18 action items on facilitating data flow and safeguarding data security.

In December 2024, the DPO released the "Data Governance Principles", outlining the guiding principles established by the Government for data governance.[12] Through data governance policies, the opening up and sharing of data with the development of analytical applications are well supported and promoted.

Digital Infrastructure for AI Development

High-performance computing is crucial for advancing medical AI research and development. The Artificial Intelligence Supercomputing Centre (AISC), established by Cyberport in Hong Kong, began its operations in December 2024 to meet the needs of local universities, Research and Development (R&D) institutes, and enterprises. By early 2026, the AISC is expected to achieve a computing power of 3,000 peta-floating point operations per second (petaFLOPS), capable of processing nearly 10 billion images per hour, positioning it as one of the largest supercomputing facilities in the region.

Encouraging Application of AI

The Government encourages B/Ds, and public organisations to leverage technologies, including AI, to enhance public services and improve internal efficiency. For instance, the Hospital Authority has implemented AI technologies to elevate its services, utilising an AI imaging model that analyses approximately 2,000 chest X-rays daily. This supports clinical decision-making and prioritises patients with higher risk factors for earlier consultation and treatment.

Conclusion

The Government strives to seize the opportunity in the digital era through the use of AI, including applications in the medical field, to enhance the quality of life for citizens. While recognising the challenges associated with AI, we have introduced guidelines like the "Ethical Artificial Intelligence Framework" and "Hong Kong Generative Artificial Intelligence Technical and Application Guideline" to provide essential guidance for B/Ds and the public in their AI initiatives. These guidelines are

---

[10] Financial Services and the Treasury Bureau of the Government of the HKSAR, "Policy Statement on Responsible Application of Artificial Intelligence in the Financial Market", 28 October 2024, https://gia.info.gov.hk/general/202410/28/P2024102800154_475819_1_1730083937115.pdf.

[11] Innovation, Technology and Industry Bureau of the Government of the HKSAR, "Home: Publications and Videos", Policy Statement on Facilitating Data Flow and Safeguarding Data Security in Hong Kong, 8 December 2023, https://www.itib.gov.hk/assets/files/Policy_Statement_Eng.pdf.

[12] Digital Policy Office of the Government of the HKSAR, "Home: Our Work: Data Governance: Enhancing Data Governance", Principles of Data Governance, 3 December 2024, https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/policy/doc/principles_of_data_governance_en.pdf.

publicly accessible through the DPO's website, encouraging various sectors to adopt it for their own governance and promoting safe and responsible AI development.

Part 1: Hong Kong SAR – 2

AI for Medical Innovation – from Medical Device's Perspective

Mr LAM Kam Chun Tommy
Senior Electronics Engineer (Medical Device)
Medical Device Division, Department of Health
The Government of the HKSAR, China

Artificial Intelligence (AI) is a rapidly advancing technology that is increasingly integrated into various aspects of human activity. Its development within the medical field is particularly noteworthy, as it aids in clinical decision-making and new drug development. In this context, this report attempts to give a brief review on the latest AI technologies, the challenges associated with their application, and the governance of AI medical device (AIMDs) in Hong Kong.

Key Capabilities of AI

The central theme of AI applications revolves around 'prediction' — the ability to understand or visualize the unknown — while aiming at achieving outcomes that surpass human capabilities.

AI's capability to abstract

In our increasingly data-driven world, the ability to distill insights from vast amounts of information has never been more crucial. AI excels at finding subtle patterns and making predictions, whether through regression analysis or more complex approaches.

The field of computer vision, particularly in medical image segmentation and related disease characterization, has greatly advanced with the introduction of Convolutional Neural Networks (CNNs) [13]. Trained with a large number of images, CNNs excel at extracting spatial and morphological features, making them invaluable for medical imaging applications such as medical image segmentation, disease characterization, histological image analysis, and examining complex structures like chromosomes. Additionally, hybrid models that combine CNNs and transformers enhance 3D medical image segmentation capabilities.

The exploration of time series signals, including biosignals such as electrocardiogram (ECG), has gained significant momentum. Models like Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) utilize memory mechanisms that enhance data interpretation. The emergence of Temporal Convolutional Networks (TCNs) has further refined our ability to uncover causal relationships, potentially leading to improved predictions of patient outcomes.

[13] Breesam, Aqeel Majeed, Sarah R. Adnan, and Shatha M. Ali. "Segmentation and Classification of Medical Images Using Artificial Intelligence: A review." *AI-Furat Journal of Innovations in Electronics and Computer Engineering* 3, no. 2 (2024): 299–320.

AI's capability to explore

The application of evolutionary algorithms and swarm intelligence in AI serves to optimize solutions in complex domains, such as drug discovery. The vast chemical space, characterized by an immense number of potential molecular combinations, poses significant challenges for traditional methodologies. However, AI's analytical capabilities are increasingly utilized to address these challenges, thereby facilitating innovative advancements in pharmaceutical development.

In radiotherapy, these algorithms optimize radiation delivery to tumors while safeguarding surrounding healthy tissues, exemplifying AI's potential in enhancing therapeutic precision. Additionally, the fusion of AI with medical treatment planning elevates our approach to healthcare interventions.

AI's capability to generate

Recent advances in AI have significantly enhanced its generative capabilities, particularly through the development of Generative Adversarial Networks (GANs) and autoencoders[14]. The advent of GANs has transformed our understanding of pattern generation. These models can create images, sounds, and even text that go beyond mere imitation. In the medical field, GANs may be employed to generate synthetic medical images, which can be used to train diagnostic models, thereby addressing the challenge of limited training samples.

On the other hand, autoencoder networks may be used to enable medical image fusion, such as combining CT and MRI scans. By learning a compressed representation of the input data and then reconstructing it, autoencoders can effectively merge information from different types of medical images, leading to new avenues for disease characterization and diagnosis.

At the forefront of generative AI is the Generative Pre-trained Transformer (GPT) model. By predicting the next word in a sequence, GPT can generate coherent and contextually relevant text. This allows it to engage in meaningful conversations and even abstract complex concepts, paving the way to more innovative medical innovations.

Challenges in AI

While AI has introduced new possibilities in the medical field, several challenges persist that hinder scalable production and deployment. Some of these challenges raise concerns for regulators.

One of the foremost issues in AI, particularly in Artificial Neural Networks (ANNs), is the balance between underfitting and overfitting. Underfitting occurs when a model is too simplistic, failing to capture the complexity of the data, resulting in poor prediction performance. Conversely, overfitting happens when a model becomes overly complex, learning noise rather than meaningful patterns,

[14] Bengesi, Staphord, Hoda El-Sayed, MD Kamruzzaman Sarker, Yao Houkpati, John Irungu, and Timothy Oladunni, "Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers", *IEEE Access*, 12 (2024), 69812–69837. https://doi.org/10.1109/access.2024.3397775.

leading to poor generalization to new data. In the AI community, various theories and practices aim to mitigate overfitting, including cross-validation, dropout, early stopping, and data augmentation. However, there are no definitive solutions, and the effectiveness of these strategies can vary depending on the specific context.

Biases are another key concern affecting performance of AI. AI systems often inherit biases present in their training data, leading to unfair and inaccurate predictions. In the aspect of medical applications, the presence of racial, gender and socioeconomic biases in training data could result in substandard clinical decisions or the perpetuation and exacerbation of healthcare disparities. Additionally, depending on applications, algorithms themselves may introduce biases, further complicating the issue.

Another major concern is hallucination, which is exclusively prominent in large language models (LLM). Hallucination can generate information that sounds credible but is factually incorrect. This lack of grounding in reality can mislead users and erode trust in AI systems. The AI community is developing alternatives to combat hallucinations, such as by using Chain-of-Thought reasoning, prompt engineering, and Retrieval Augmented Generation, with a view to reducing the possibilities of hallucinations.

Another challenge is the lack of explainability in many AI models, often described as "black boxes". This obscurity makes it difficult for users to understand the decision-making process. This opacity raises accountability concerns, particularly in critical areas like healthcare. Explainable AI (xAI) has emerged as a pivotal topic in research, with organizations such as the Institute of Electrical and Electronics Engineers (IEEE) working towards developing well-defined harmonized standards to enhance transparency in AI systems[15].

Ethical concerns are paramount, especially regarding **privacy**. The data needed for AI often includes sensitive personal information, raising significant privacy issues. We must be vigilant about **cybersecurity vulnerabilities**. AI systems can be susceptible to adversarial attacks and data poisoning, compromising their reliability and safety.

Local Regulatory Framework for Medical Devices

The integration of AI technology into the medical field is becoming increasingly significant. Software as Medical Devices (SaMD) and Artificial Intelligence Medical Devices (AIMD) have emerged prominently in recent years, particularly in biomedical imaging analysis, and are increasingly making their mark in biosignal analysis and diagnostics.

In Hong Kong, there is currently no overarching legislation governing the manufacture, import, distribution, supply, and use of medical devices (MDs). However, depending on their nature and characteristics, some may be regulated by existing legislation.[16]

[15] IEEE Computational Intelligence Society/ Standards Committee (CIS/SC), "eXplainable AI Working Group", Accessed 19 March 2025, https://sagroups.ieee.org/2976/.
[16] Examples include the Pharmacy and Poisons Ordinance (Cap. 138), the Radiation Ordinance (Cap. 303), and the Telecommunications Ordinance (Cap. 106).

To establish a long-term statutory regulatory framework for the safety, quality, and performance of medical devices, the Medical Device Division (MDD) of the Department of Health introduced the voluntary[17] Medical Device Administrative Control System (MDACS) in 2004[18]. This framework combines a premarket listing approach with post-market controls, aimed at enhancing the accessibility of new medical technologies while carefully assessing associated risks, all in the interest of safeguarding public health. The MDACS is regularly reviewed, updated and aligned with international standards, drawing on recommendations from the International Medical Device Regulators Forum (IMDRF) and the Global Harmonization Working Party (GHWP).

Technical Reference on AIMD

In response to the growing trend of AI adoption in medical devices, the MDD published a technical reference on AIMDs in 2024[19]. This document, which applies to all AIMDs[20] within the scope of the MDACS and forms part of the MDACS listing requirements, sets out the compliance requirements under the MDACS, including dataset application, AI model selection, performance evaluation, and post-market monitoring protocols.

Key considerations during evaluation of AIMDs include whether datasets used to train models are representative of the patient demographics, such as race, age, and gender, so as to prevent bias and ensure adequate level of generalization. Additionally, the type of AI model and the associated clinical evaluations will be assessed. For AIMDs with continuous learning capabilities, further information on how these devices will learn from real-world data will also be evaluated. Relevant safety mechanisms and software version control must be implemented to mitigate the impact of anomalies and facilitate reverting to previous algorithm versions when necessary.

Regarding data privacy and ethical concerns surrounding AI, including medical AI, relevant local responsible persons for the medical device shall also be aware of other applicable requirements, frameworks and practices in respect of AI applications, such as those issued by the Digital Policy Office and the Office of the Privacy Commissioner for Personal Data.

Developing Hong Kong into a Medical Innovation Hub

As part of the initiatives of 2023 Policy Address to develop Hong Kong into a Health and Medical Innovation Hub and to establish an authority that registers medical products under the "primary evaluation" approach, the Department of Health has officially established the **Preparatory Office for the Hong Kong Centre for Medical Products Regulation (CMPR)** on 5 June 2024[21]. The

[17] Prior to the establishment of a statutory regulatory framework, listing and application for listing, whether for medical devices or as traders (i.e. importers, distributors, and local manufacturers) under the MDACS, are voluntary.
[18] Medical Device Division of the Department of Health of the Government of the Hong Kong Special Administrative Region of the People's Republic of China ("The Government of the HKSAR"), "Medical Device Administrative Control System", Accessed 19 March 2025, https://www.mdd.gov.hk/en/mdacs/index.html.
[19] Medical Device Division of the Department of Health of the Government of the HKSAR, "Artificial Intelligence Medical Devices (AI-MD) Technical Reference: TR-008", Revised on 15 November 2024, https://www.mdd.gov.hk/filemanager/common/mdacs/TR008.pdf.
[20] The Technical Reference is not compulsory for any AIMD that is not listed in the MDACS.
[21] The Government of the HKSAR, "DH Establishes Preparatory Office for Hong Kong Centre for Medical Products Regulation", 5 June 2024, https://www.info.gov.hk/gia/general/202406/05/P2024060500165.htm.

Preparatory Office will put forward proposals and steps for the formal establishment of the CMPR, and to study the potential restructuring and strengthening of the regulatory and approval regimes for medical products.

In view of the fact that the CMPR is targeted to be established in a few years' time, the MDD and Preparatory Office of CMPR will keep abreast of the latest knowledge and developments in the medical device and pharmaceutical field, including the application and relevant governance of AI, as well as considering what appropriate and pragmatic guidance should be derived under the CMPR, with a view to discharging its regulatory functions efficiently and effectively.

In light of the rapid advancements in AI, it is crucial for regulatory frameworks to be adaptive and flexible. This may necessitate the adoption of a risk-based approach that allows for adjustments of regulatory regime as technologies progress. Specifically, in the realm of medical devices, the evolving landscape of AI in medical technology and the corresponding international regulatory environment will be thoroughly considered during the transition from voluntary MDACS listing to future statutory control.

Conclusion

The potential of AI to revolutionize healthcare is significant; however, it necessitates a holistic and agile approach to managing these technologies. The MDD recognizes the growing trend of AI adoption in healthcare and is actively working to strike a balance between technological advancement and the governance of AIMDs. Through collaboration with stakeholders, we can ensure that AI serves as a positive force, enhancing care quality and improving health outcomes for all.

Part 1: Hong Kong SAR – 3

A Trustworthy Regulatory Environment for AI-enabled Assisted Reproduction in Hong Kong

Prof Calvin Wai-Loon Ho JSD PCLT FRSPH
Associate Professor, Faculty of Law, Monash University, Melbourne, Australia;
Research Fellow, Centre for Medical Ethics & Law, University of Hong Kong,
Hong Kong SAR, China

(The author thanks Ms Cordelia Chan for preparing a draft summary of the presentation, and Dr Karel Caals for his comments.)

I. Introduction

Digital technology is widely used in healthcare,[22] and assisted reproduction (AR) is no exception. Notably, artificial intelligence (AI) is now used in embryo selection to increase the success of *in vitro fertilization* (IVF). However, there is (at least at the time of writing) limited evidence to show that such use of AI greatly advances clinical outcome. A recent study (to be discussed below) appears to point to the contrary, even if there may be savings in labour cost. In the context of IVF,[23] I highlight the need to establish a regulatory environment that facilitates evidence generation for AI-based medical interventions in Hong Kong.[24] Currently, it can be argued that the regulatory regime in Hong Kong does not adequately support the evaluation of innovation in AI-enabled AR.

In Part II of this summary, I broadly consider the regulatory regime for AR in Hong Kong and its focus on sufficiency of information for the purposes of informed consent. I then consider in Part III, why a focus on informed consent has not been adequate in relation to novel reproductive technologies like non-invasive prenatal testing (NIPT). Recent findings from a clinical trial of an AI-enabled IVF treatment are broadly discussed in Part IV, and its regulatory implications are then considered in Part V.

II. Regulatory Emphasis on Patient Choice

In Hong Kong SAR, IVF is governed by the Human Reproductive Technology Ordinance (HRTO).[25] The HRTO regulates the provision of reproductive technology procedures; the conduct of embryo research; the handling, storing or disposing of gametes or embryos used or intended for use in connection with reproductive technology procedures (RTP) or embryo research; and surrogacy arrangement. The HRTO establishes two important regimes: Firstly, it sets out a licensing scheme governing the assisted reproductive technology (ART) clinics under the Human Reproductive

[22] Mauro, Marianna, Guido Noto, Anna Prenestini, Fabrizia Sarto, "Digital Transformation in Healthcare: Assessing the Role of Digital Technologies for Managerial Support Processes", *Technological Forecasting and Social Change* 209 (2024):123781. https://doi.org/10.1016/j.techfore.2024.123781; See also: Craig, Adam T, Lawford Harriet, Miller Maggie, Chen-Cao Liuyi, Woods Leanna, Liaw Siaw-Teng, et al., "Use and Impact of Digital Technology in Supporting Health Providers Deliver Care in Low- and Low-middle-income Countries: A Systematic Review Protocol." *PLoS ONE* 20, no. 2 (2025): e0319190. https://doi.org/10.1371/journal.pone.0319190.

[23] For an overview of the regulation of medical AI in the Greater Bay Area, see: Ho, Calvin Wai Loon. "Convergence in the Regulation of Artificial Intelligence Software as Medical Device in the Guangdong-Hong Kong-Macao Greater Bay Area of China. in *Research Handbook on Health, AI and the Law,* edited by Barry Solaiman, & I. Glenn Cohen, (1st ed., pp. 355–372). (Cheltenham: Edward Elgar, 2024). https://doi.org/10.4337/9781802205657.ch20

[24] Prof Li Du discussed a number of regulatory issues regarding the application of AI in assisted reproduction in Mainland China and Macau SAR during his presentation at this conference. Please refer to his expanded abstract in this conference report.

[25] Human Reproductive Technology Ordinance, Cap.561, Laws of Hong Kong.

Technology (Licensing) Regulation.[26] Secondly, it provides specific guidance under the *Code of Practice on Reproductive Technology and Embryo Research*[27] and *Licensing Manual for Reproductive Technology Centres*.[28] This is implemented by the Council on Human Reproductive Technology, which itself is established under section 4 of the HRTO. The Council is funded through fees payable under the Human Reproductive Technology (Fees) Regulation.[29]

The regulatory approach in Hong Kong is similar to those of England and the United States in terms of its emphasis on informed consent.[30] It prioritises the provision of sufficient information to patients so that they may make the right decisions about their treatment and care. This can be seen, for instance, in the informational requirements set out in the *Code of Practice*,[31] as well as in professional governance.[32]

III. NIPT for Down Syndrome

Despite regulatory safeguards, sufficiency of information may not necessarily provide assurance as to the adequacy of the medical intervention, particularly where evidence of effectiveness is scant owing in part to its novelty. Consider NIPT for Down Syndrome, for instance. Research indicates that women were not properly supported in decision-making. Olivia Ngan *et al*. observed the following issues:[33]

(i) Obstetric providers have reportedly perceived less need for consent procedures for NIPT compared to invasive prenatal diagnosis (IPD);

(ii) Reliance on general information pamphlets;

(iii) Otherwise, discussion tended to focus on termination of pregnancy;

(iv) Lack of clarity on referral between public and private sectors, and the attending responsibilities of healthcare professionals;

(v) Out-of-pocket payment that exacerbates unequal access;

(vi) Since obstetricians are responsible for informed consent, time constraints is a challenge, considering the steps involved (pre-test and post-test counselling and follow-up) and the complexity of newer modalities like expanded NIPT (eNIPT) and the prospect of incidental findings that it raises; and

---

[26] Human Reproductive Technology (Licensing) Regulation, Cap.561A, Laws of Hong Kong.
[27] Council on Human Reproductive Technology, "Code of Practice on Reproductive Technology and Embryo Research", January 2024, https://www.chrt.org.hk/english/publications/files/code2024.pdf.
[28] Council on Human Reproductive Technology, "Licensing Manual for Reproductive Technology Centres", April 2025, https://www.chrt.org.hk/english/service/files/MCL.pdf.
[29] Human Reproductive Technology (Fees) Regulation, Cap.561B, Laws of Hong Kong.
[30] Perrot, Adeline, Horn, Ruth. "The Ethical Landscape(s) of Non-invasive Prenatal Testing in England, France and Germany: Findings from a Comparative Literature Review." *European Journal of Human Genetics* 30 (2022): 676–681. https://doi.org/10.1038/s41431-021-00970-2.
[31] Council on Human Reproductive Technology, "Code of Practice on Reproductive Technology and Embryo Research", January 2024, https://www.chrt.org.hk/english/publications/files/code2024.pdf.
[32] The Medical Registration Ordinance, Cap.161, empowers the Medical Council of Hong Kong (MCHK) to regulate professional conduct of registered medical practitioners in the territory. The Code of Professional Conduct of the MCHK highlights the importance of a patient's informed consent in clinical care. See: Medical Council of Hong Kong, "Code of Professional Conduct", October 2022 (revised), https://www.mchk.org.hk/english/code/files/Code_of_Professional_Conduct_(English_Version)_(Revised_in_October_2022).pdf
[33] Ngan, Olivia Miu Yung, Huso Yi, Shenaz Ahmed. "Service Provision of Non-invasive Prenatal Testing for Down Syndrome in Public and Private Healthcare Sectors: A Qualitative Study with Obstetric Providers." *BMC Health Services Research* 18, no. 1 (2018): 731, http://doi.org/10.1186/s12913-018-3540-9.

(vii) Challenge of "non-directive" counselling.

In the light of the above, the sufficiency and adequacy of guidance provided to patients will require careful consideration when scaling up the application of NIPT to eNIPT.[34] Even if sufficient information is available, it may not be adequate if patients do not understand the value of the intervention, or if they are unable to benefit fully from it. We also need to consider how to safeguard against the "hype" around novel technology, to ensure that patients clearly understand what they are ultimately paying for, and that this payment is fair in terms of its expected outcomes. These assurances help to build and sustain trust between healthcare provider and the care recipient, and more broadly within the health system. Additionally, it helps to avoid the kind of regulatory backlash that was witnessed when NIPT was initially introduced in mainland China.[35]

NIPT was introduced into clinical practice on mainland China and in the Hong Kong SAR at around 2010 to 2011.[36] The hype around this (then) new technology propelled its uptake, but concerns about the technology's reliability and possible harms triggered strong regulatory responses. Between 2011 to 2014, several Chinese "cell-free" DNA-based NIPT tests were available, but access was limited by their relatively high cost.[37] In 2014, Chinese regulators suspended all prenatal genetic testing (including NIPT) until the implementation of new regulations, due to concerns about the highly variable quality of the tests and unsubstantiated claims by commercial providers. Since then, conditional marketing permits have been granted for a small number of tests developed by well-known manufacturers, which are mainly available in the private sector. Some Chinese provinces have also included NIPT for selected indications in their state-sponsored parental care, where partial reimbursement of the cost is provided. It is on this regulated basis that evidence generation has been enabled.[38]

IV. AI-enabled IVF and evidence generation

Prior regulatory experience with reproductive technologies like NIPT and cfDNA tests offers insights as to anticipatory regulatory measures that may be considered as AI modalities are being incorporated into assisted reproduction. Where evidence generation is concerned, Hong Kong does not currently have a regulatory framework that supports the production of reliable evidence to guide policy and patient decision-making, apart from generic clinical trial regulatory guidelines.[39] In the context of assisted reproduction, IVF protocols are complex and require intensive monitoring, and different decision responsibilities are shared between clinicians and embryologists. Decision making happens at different stages of embryonic development before implantation to ensure a successful pregnancy. Several of these decisions have a solid evidence base, but many are highly subjective and will vary greatly depending on clinical experience. Data-driven approaches and AI technology are now being

[34] Taylor-Sands, Michelle, Molly Johnston, Catherine Mills. "Should the Scope of NIPT be Limited by A 'Threshold of Seriousness'?" *European Journal of Human Genetics* 33, no. 2 (2025):189–193. http://doi.org/10.1038/s41431-024-01684-x.

[35] Mei, Lin, Qi Tang, Baiyu Sun and Lingzhong Xu, "Noninvasive Prenatal Testing in China: Future Detection of Rare Genetic Diseases?" *Intractable & Rare Diseases Research* 3, no.3 (2024): 87–90. http://doi.org/10.5582/irdr.2014.01012.

[36] Allyse, Megan, Minear Mollie A., Berson Elisa, Sridhar Shilpa, Rote Margaret, Hung Anthony and Chandrasekharan Subhashini. "Non-invasive Prenatal Testing: A Review of International Implementation and Challenges." *International Journal of Women's Health* 7 (2015): 113–26. http://doi.org/10.2147/IJWH.S67124.

[37] Rafi, Imran, Melissa Hill, Judith Hayward and Lyn S Chitty. "Non-invasive Prenatal Testing: Use of Cell-free Fetal DNA in Down Syndrome Screening." *British Journal of General Practice* 67, no. 660 (2017): 298–299. http://doi.org/10.3399/bjgp17X691625.

[38] Chan, Tian, Deng, Tao, Zhu, Xiuhuang. et al., "Evidence of Compliance with and Effectiveness of Guidelines for Noninvasive Prenatal Testing in China: A Retrospective Study of 189,809 Cases." *Science China Life Sciences* 63 (2020): 319–328. https://doi.org/10.1007/s11427-019-9600-0; See also: Shang, Wenru, Wan Yang, Chen Jianan, et al., "Introducing the Non-invasive Prenatal Testing for Detection of Down Syndrome in China: A Cost-Effectiveness Analysis." *BMJ Open* 11 (2021): e046582. http://doi.org/10.1136/bmjopen-2020-046582.

[39] See Section 36B of the Pharmacy and Poisons Regulations (Cap. 138A).

incorporated to facilitate objective, consistent and optimal decision-making, and to drive individualised treatment. Such approaches range from algorithmic drug dosing tools to 'human-in-the-loop' AI clinical decision support for embryo selection. These AI models are developed based on about 1 million cycles undertaken annually worldwide.

The conventional approach for embryo selection requires daily morphological assessment of the embryo, which is very labour-intensive. A clinical trial was done recently to compare the conventional approach and the AI-enabled approach to embryo selection, where all embryos were cultured in a time-lapsed incubator. On day 5, patients were randomly assigned to have their embryos selected by conventional morphological methods or by a deep learning algorithm developed for the AI-enabled approach. In this study, non-inferiority could not be demonstrated; it found clinical pregnancy rates of 48.2% for conventional methods and 46.5% for deep learning algorithms. However, the deep learning approach reduced evaluation time in the laboratory and was less labour-intensive compared to simpler machine learning methods, since embryologists did not have to check and correct automatic annotations of important cell-cycle parameters (which is laborious). However, it incurred a much higher net cost: the equipment cost a lot more than conventional incubators, even though labour cost decreased and laboratory efficiency increased.[40]

This means that patients would be paying a much higher price for almost no improvement in outcome. The question therefore becomes one of "how well is the patient informed?". For many of these treatment add-ons, there is limited evidence as to their effectiveness. However, given the "hype" around these new AI-enabled approaches, patients are prepared to pay more for them. We therefore turn to consider the non-neutral effects of IVF treatment "add-ons". These are non-essential extra treatment which have either no or limited evidence (typically from low quality retrospective studies) of improving clinical outcomes, but which imposes substantial additional costs. The question of who protects patients and the healthcare system from exploitation that is linked to the "hype" or false hope around new technologies also comes up: whether it is a case of *caveat emptor*, or if patients are not adequately informed, or if it is a matter to be addressed as an issue of medical professionalism. There are at least two broader issues about the regulatory status quo in Hong Kong that warrant further consideration. First, it may be asked if a 'light-touch' approach to AI-enabled ART services advances public good, particularly in terms of how resources are used. Second, the impact on public trust in the medical profession and on the integrity of the health data ecosystem are also related queries that arise.

V. An Enabling Regulatory Environment

It is beyond the scope of this short presentation to engage fully with these questions and challenges. However, the proposition that I advance here and elsewhere is the need for AI-enhanced healthcare interventions to be nestled within a supportive regulatory framework or infrastructure that has the following characteristics:

(i) Enables participation within the existing data and device-based ecosystems;[41]

---

[40] Illingworth, Peter J., Christos Venetis, David K. Gardner, Scott M. Nelson, Jørgen Berntsen, Mark G. Larman, Franca Agresta, Saran Ahitan, Aisling Ahlström, Fleur Cattrall, Simon Cooke, Kristy Demmers, Anette Gabrielsen, Johnny Hindkjær, Rebecca L. Kelley, Charlotte Knight, Lisa Lee, Robert Lahoud, Manveen Mangat, Hannah ParkAnthony Price, Geoffrey Trew, Bettina Troest, Anna Vincent, Susanne Wennerström, Lyndsey Zujovic, Thorir Hardarson, "Deep Learning Versus Manual Morphology-Based Embryo Selection in IVF: A Randomized, Double-Blind Noninferiority Trial." *Nature Medicine* 30, no. 11 (2024): 3114-3120, http://doi.org/10.1038/s41591-024-03166-5.

[41] Ho, Calvin Wai Loon. "Implementing the Human Right to Science in the Regulatory Governance of Artificial Intelligence in Healthcare." *Journal of Law and the Biosciences* 10, no. 2 (2023), https://doi.org/10.1093/jlb/lsad026. See also: Ho, Calvin Wai Loon, Ali, Joseph and Caals, Karel. "Ensuring Trustworthy Use of Artificial Intelligence and Big Data Analytics in Health Insurance." *Bulletin of the World Health Organization* 98, no. 4: 263–269. https://doi.org/10.2471/BLT.19.234732.

(ii) Supports role-based collaboration and thereby also give meaningful effect to the "human-in-the-loop" approach;[42]

(iii) Enables ethics governance to facilitate (i) and (ii) above, while ensuring that the safety and wellbeing of human participants are secure;[43]

(iv) Shifts evaluation of trustworthiness from transactional to environmental;[44] and

(v) Capitalises on epistemic hybridisation (particularly the normative domains of ethics, law and human rights).[45]

Implicit to these characteristics is the need to continuously monitor, evaluate and refine the effectiveness of AI modalities that are incorporated into clinical care. There are no clear markers, but some relevant measures in the ART context to evaluate evidence could be increased pregnancy rate, improved laboratory work, reduced inefficiency, as well as verified safety of various components. However, questions remain as to how one factor weighs against the other factors, and who should have the responsibility to decide on the most relevant and appropriate outcome measures.

While there is a clear need to support the real-world testing of AI applications, there is still a need to sustain the distinction between research and therapy, enabling evidence generation on different fronts. To be sure, this does not mean that we should forego retrospective data analysis, assuming the relevant data is available.[46] Instead, we should think about how research and therapy can be integrated while supporting randomised controlled trials (RCTs), as an important source of evidence. We also need to rethink participation, particularly in the IVF space. This is consistent with the normative commitments in the International Covenant on Economic, Social and Cultural Rights, which are applicable in Hong Kong.[47]

---

[42] Ho, Calvin Wai Loon. "Securing the 'Human' in the Generalization of Risk Stratification Algorithms through the Human Right to Science." in *Promoting the "Human" in Law, Policy, and Medicine: Essays in Honour of Bartha Maria Knoppers,* edited by Edward S. Dove, Vasiliki Rahimzadeh and Michael J. S. Beauvais, (1st ed., pp. 342–362). (Global Health, Human Rights and Social Justice; Vol. 1). (Germany: Brill, 2025), https://doi.org/10.1163/9789004688544_048.

[43] Ho, Calvin Wai Loon and Caals, Karel. "A Call for an Ethics and Governance Action Plan to Harness the Power of Artificial Intelligence and Digitalization in Nephrology." *Seminars in Nephrology* 41, no. 3 (2021): 282–293. https://doi.org/10.1016/j.semnephrol.2021.05.009.

[44] Ho, Calvin Wai Loon and Caals, Karel. "How the EU AI Act Seeks to Establish an Epistemic Environment of Trust." *Asian Bioethics Review* 16 (2024): 345–372. https://doi.org/10.1007/s41649-024-00304-6.

[45] Ho, Calvin Wai Loon. "When Learning is Continuous: Bridging the Research–Therapy Divide in the Regulatory Governance of Artificial Intelligence as Medical Devices." in *The Cambridge Handbook of Health Research Regulation,* edited by Graeme Laurie, Edward Dove, Agomoni Ganguli-Mitra, Catriona McMillan, Emily Postan, Nayha Sethi and Annie Sorbie, (1st ed., pp. 277–286). (Cambridge: Cambridge University Press, 2021), https://doi.org/10.1017/9781108620024.035

[46] Prof Vera Raposo, in her presentation at this conference, discussed the phenomenon that the relevant data might be in private hands and, therefore, not available. This conference report does not contain any expanded abstract, summary or transcript of her presentation.

[47] Department of Justice of the Government of Hong Kong SAR. "Application of the International Covenant on Economic, Social and Cultural Rights in Hong Kong." *Basic Law Bulletin* 17 (2015): 3–10, https://www.doj.gov.hk/en/publications/pdf/basiclaw/basic17_3.pdf

VI. Conclusion

In the light of recent advances in reproductive technologies, this paper highlights the need to establish a regulatory environment that facilitates evidence generation in Hong Kong. In support of this proposition, previous regulatory experiences with novel reproductive technologies like NIPT and related modalities are discussed, along with recent findings from a clinical trial of an AI-enabled IVF treatment. General features of this proposed regulatory environment, all of which consistent with normative commitments of Hong Kong SAR, are also advanced in this paper.

Part 2: Macau SAR and Mainland China – 1

Application of AI in Assisted Reproduction: Regulatory Issues in Mainland China and Macau

Mr Zhangyu Wang* and Prof Li Du**

* PhD Student, Faculty of Law, University of Macau, Macau SAR, China
** Associate Professor, Faculty of Law, University of Macau, Macau SAR, China

I. Introduction

In light of decreasing fertility rates globally, assisted reproduction technologies (ART), particularly in-vitro fertilization (IVF), have played a crucial role in treating human infertility and ensuring reproductive autonomy. Most recently, advances in information and communication technologies have introduced digital tools and artificial intelligence (AI) into ART treatments. Compared to conventional approaches, where effectiveness is often heavily reliant on a practitioner's experience and expertise, AI-enhanced ART could lead to more accurate assessments of clinical indicators and medical images. This could significantly reduce both errors and the subjectivity of visual inspections, thereby mitigating the risk of misrepresentation by practitioners with limited expertise or inadequate training.[48] For instance, the Israeli company Fairtility has claimed that their AI-powered product can accurately select IVF embryos with the likelihood of 78% successful implantation, thus outperforming the average 60% accuracy rate of embryologists.[49]

Although AI holds the potential to enhance the success rate of IVF, research of AI-enabled ART and its adoption in clinical settings still face significant ethical and legal hurdles.[50] For instance, research and development of AI-enabled ART systems requires a large amount of high-quality health data to minimize the risks of inaccuracy and reduce the predictive and diagnostic biases of AI systems, which jeopardize the quality of clinical applications.[51] In this regard, legal regimes that support the use and smooth cross-border flow of health-related personal data, especially of medical records and genetic data, can significantly enhance research on AI-enabled ART and its utility for diverse ethnic and racial populations. In our presentation, we explored the relevant laws of Mainland China and Macau concerning the development of AI-enabled ART and identified key regulatory challenges that could influence the progress and mutual collaboration in this emerging clinical field in Mainland China, Macau, and internationally.

---

[48] Wang, Guangyu et al., "A Generalized AI System for Human Embryo Selection Covering the Entire IVF Cycle via Multi-modal Contrastive Learning," *Patterns* 5, no.7 (2024): 100985, https://doi.org/10.1016/j.patter.2024.100985; Zaninovic, Nikica and Zev Rosenwaks, "Artificial Intelligence in Human In Vitro Fertilization and Embryology," *Fertility and Sterility* 114, no. 5 (2020): 914-20, https://doi.org/10.1016/j.fertnstert.2020.09.157.

[49] Jeffay, Nathan. "Israeli Embryo-selecting System to Boost IVF Success Gets European Green Light," *The Times of Israel* (2022), accessed January 14, 2025, https://www.timesofisrael.com/israeli-embryo-selecting-system-to-boost-ivf-success-gets-european-green-light/.

[50] Tamir, Sivan. "Artificial Intelligence in Human Reproduction: Charting the Ethical Debate over AI in IVF," *AI and Ethics* 3, (2023): 947–61, https://doi.org/10.1007/s43681-022-00216-x; Zammit, Raymond. "Ethical Issues of Artificial Intelligence & Assisted Reproductive Technologies," *International Journal of Prenatal & Life Sciences* (2023), https://doi.org/10.24946/IJPLS/2023132305.

[51] Norori, Natalia et al., "Addressing Bias in Big Data and AI for Health Care: A Call for Open Science," *Patterns* 2, no. 10 (2021): 100347, https://doi.org/10.1016/j.patter.2021.100347.

II. Regulatory challenges in Mainland China

In Mainland China, the regulation of the use and cross-border transfer of health data encompasses three major legal areas: namely, the general personal data protection law, the law and regulations on the management of human genetic resources, and the medical record management regulations. Based on our previous studies,[52] we identified several significant regulatory challenges related to this complex regulatory environment that could hinder research into AI-enabled ART.

i. Strict informed consent requirement

Chinese personal data protection laws require that informed consent is obtained for the use or transfer of personal health information of patients and that it is reobtained when the purpose or method of processing personal information or the type of personal information to be processed changes.[53] Meanwhile, the current regulatory framework has not yet established specific exemption or facilitation measures that would enable scientists to bypass "repeated informed-consent process that…may dilute the purpose of research and what it can achieve while bringing unnecessary disturbances to participants."[54] In this regard, the strict rules for informed consent increase the compliance cost of using AI in ART procedures and may even render compliance impossible.

ii. Rigorous management of human genetic data cross-border transfers

Currently, the *Biosecurity Law* and *Administrative Regulations on Human Genetic Resources* (HGR Regulation) serve as the primary legislation and regulation governing the use and cross-border transmission of human genetic data of China.[55] The *Biosecurity Law* and HGR Regulation, among other things, prohibit foreign entities from, within China's territory, collecting or preserving human genetic data of China and prohibit foreign entities from providing human genetic data of China to out of China's territory, with retaining limited permission for outbound cross-border flow of human genetic data under specific circumstances and under strict administrative compliance procedures. The *Biosecurity Law* and HGR Regulation impose requirements on both (i) domestic data owners' provision or open access of human genetic data to foreign data users (for which prior reporting to and submitting of data backup to National Health Commission (NHC) is necessary) and (ii) the use of human genetic data of China in international scientific research collaboration (for which approval by

---

[52] Wang, Zhangyu, Benjamin Gregg, and Li Du, "Regulatory Barriers to US-China Collaboration for Generative AI Development in Genomic Research," *Cell Genomics* 4, no.6 (2024): 100564, https://doi.org/10.1016/j.xgen.2024.100564.

[53] For example, according to Article 14 of the Personal Information Protection Law of People's Republic of China (中華人民共和國個人信息保護法) ("PIPL"), where the processing of personal information is based on the consent of the individual concerned, such consent shall be given by the individual concerned in a voluntary and explicit manner under the condition of full knowledge; where laws and administrative regulations provide that the processing of personal information shall be subject to the separate consent or written consent of the individual concerned, such provisions shall prevail; where the purpose or method of processing personal information or the type of personal information to be processed changes, the consent of the individual concerned shall be obtained again. Please also refer to Articles 29 and 39 of the PIPL and Article 9 of the Detailed Rules for the Implementation of Administrative Regulations on Human Genetic Resources (Ministry of Science and Technology of the People's Republic of China, "Detailed Rules for the Implementation of Administrative Regulations on Human Genetic Resources (人類遺傳資源管理條例實施細則)," 2023).

[54] Li, Xiaojie, Yali Cong, and Ruishuang Liu. "Research under China's Personal Information Law," *Science* 378, no. 6621 (2022): 713-15, https://doi.org/10.1126/science.abq7402.

[55] State Council of the People's Republic of China, "Administrative Regulations on Human Genetic Resources of the People's Republic of China (中華人民共和國人類遺傳資源管理條例)," 2019 (amended in 2024), Articles 1-2 and 7; Standing Committee of the National People's Congress, "Biosecurity Law of the People's Republic of China (中華人民共和國生物安全法)," 2020 (amended in 2024), Chapter 6.

NHC is necessary).[56] In both cases, NHC exerts a strict control and the act is impermissible if it will endanger public health, national security or social public interest, and, if the provision or open access in (i) might affect public health, national security or social public interest, passing of a specific security review is required.[57]

In addition, the exportation of personal data (including but not limited to sensitive personal data such as genetic data) will trigger security evaluation by the Cyberspace Administration of China (CAC), if the intended data exporter has exported an aggregate amount of personal data that exceeds the legally stipulated numerical threshold applicable or if the intended data exporter is a critical information infrastructure operator or if the data to be exported constitutes important data.[58] Consequently, China's stringent genetic data protection regulations make cross-border transmission of genetic data difficult, with the situation being exacerbated by cumbersome and non-transparent administrative approval procedures, which could significantly undermine cross-border collaborative research on AI-enabled ART treatments.[59]

iii. Localization requirements for health data generated in medical sector

Legal requirements wielding rigorous management of medical records in China may become another layer of constraint to data use for AI training. According to *Regulations for Medical Institutions on Medical Records Management*, except for medical personnel who provide diagnostic and treatment services to patients and departments/ personnel authorised by the relevant authorities or medical institutions as more particularly described in those regulations, other medical institutions and personnel who intend to access or borrow medical records for scientific research and teaching purposes shall apply to the medical institution where the patient is treated.[60] Approved access to medical records should be obtained onsite and the relevant medical records should be returned immediately afterwards, while the borrowed medical records should be returned within three working days.[61] China's NHC further clarified in a policy document released in 2018 that "… patient information, diagnosis and treatment data … are strictly managed to protect patients' privacy and information security. Patient information and other sensitive data should be stored within the country."[62] Therefore, any access and borrowing can only be done within China's borders or through

---

[56] State Council of the People's Republic of China, "Administrative Regulations on Human Genetic Resources of the People's Republic of China (中華人民共和國人類遺傳資源管理條例)," 2019 (amended in 2024), Articles 22 and 28; Standing Committee of the National People's Congress, "Biosecurity Law of the People's Republic of China (中華人民共和國生物安全法)," 2020 (amended in 2024), Articles 55, 56 and 57.

[57] *Ibid.*

[58] Standing Committee of the National People's Congress, "Personal Information Protection Law of the People's Republic of China (中華人民共和國個人信息保護法)," 2021, Article 38; Cyberspace Administration of China, "Provisions on Promoting and Regulating Cross-Border Data Flows (促進和規範數據跨境流動規定)," 2024, Article 7 and 8.

[59] Wang, Zhangyu, Benjamin Gregg, and Li Du, "Regulatory Barriers to US-China Collaboration for Generative AI Development in Genomic Research" *Cell Genomics* 4, no.6 (2024): 100564, https://doi.org/10.1016/j.xgen.2024.100564; Chen, Yongxi and Lingqiao Song, "China: Concurring Regulation of Cross-border Genomic Data Sharing for Statist Control and Individual Protection," *Human Genetics* 137, no.8 (2018): 605-15, https://doi.org/10.1007/s00439-018-1903-2.

[60] State Administration of Traditional Chinese Medicine, and National Health Commission (National Health and Family Planning Commission), "Regulations for Medical Institutions on Medical Records Management (醫療機構病歷管理規定)," 2013, Article 16.

[61] *Ibid.*

[62] National Health Commission, "Notice on Further Promoting the Informatization of Medical Institutions with Electronic Medical Records as the Focal Point (關於進一步推進以電子病歷為核心的醫療機構信息化建設工作的通知)," (2018), paragraph 6.

servers based in China, meaning that the use of medical records in cross-border research collaborations seems less attainable.

In addition, the National Health and Family Planning Commission (a former agency of NHC) enacted the *Measures for the Administration of Population Health Information* in 2014, designating the scope of Population Health Information (PHI) to include basic population information, medical and health service information, and other information generated in the process of service and administration by medical institutions at all levels within China.[63] The Measures has unequivocally prohibited medical institutions from storing PHI in servers outside China, and from hosting or renting servers outside the territory of China.[64] Under the localization provision, any information generated in the medical sector may become PHI and thus is less likely to be used for cross-border collaboration to develop AI-enabled ART treatments.

III. Regulatory challenges in Macau

In Macau, the use of medical data in AI-assisted ART is governed by three primary laws: the 39/99 *Macau Civil Code*, the 14/2023 *Medical Assisted Reproduction Technology Law*, and the 8/2005 *Personal Data Protection Act*. Pursuant to Article 42 of the 14/2023 *Medical Assisted Reproduction Technology Law* (MARTL), on the condition that the applicability of the "several following" articles of the MARTL are not affected, the relevant provisions of the personal data protection law shall apply to the personal data relating to the medical assisted reproductive procedures, donors, beneficiaries and children born. Article 45 of the MARTL provides that, without prejudice to Articles 26 and 44.2, personal data relating to medical assisted reproduction shall be accessed for medical purposes only and that access, for the purposes of medical research, of any personal data that could directly or indirectly identify any person involved is prohibited except where the data subject gives written express consent. Article 44.1 of the MARTL provides that, without prejudice to Article 26, only public hospital's or medical assisted reproduction unit's technical head or medical personnel specified thereby are allowed to access personal data relating to medical assisted reproduction after the end of clinical use whereas Article 44.2 of the MARTL provides that auditors and supervisors are allowed to access personal data relating to medical assisted reproduction within the scope referred to in Article 18. The *Personal Data Protection Act* has established three mechanisms to facilitate the use and transfer of personal data. For example, the obligation to provide the data subject with the information specified in Article 10 of the Act could potentially be waived for the processing of personal data for the purposes of historical or scientific research when the provision of such information is either impossible or requires a disproportionate level of effort.[65] Moreover, the processing of personal data related to an individual's health, sex life, and genetic information, could be allowed if necessary for preventive medicine, medical diagnosis, health care provision, treatment, or management of health care services, provided that such data are processed by a health professional bound by an obligation of secrecy or by another person similarly subject to an occupational obligation of secrecy, and it is notified to the public authority and where suitable safeguards are provided.[66] Additionally, in cases of cross-border data transfer to a destination in which the legal system does not ensure an adequate

[63] National Health Commission (National Health and Family Planning Commission), "The Measures for the Administration of Population Health Information (for Trial Implementation) (人口健康信息管理辦法試行)," 2014, Article 3.

[64] *Ibid*., Article 10.

[65] Legislative Assembly of Macau SAR., "Personal Data Protection Act (Law No. 8/2005) 澳門特別行政區 第 8/2005 號法律 個人資料保護法," 2005, Article 10.5(3).

[66] *Ibid.*, Article 7.4. Please also refer to the requirements imposed in respect of personal data by the 14/2023 Medical Assisted Reproduction Technology Law (Legislative Assembly of Macau SAR., "Medical Assisted Reproduction Technology Law (Law No. 14/2023) 澳門特別行政區 第 14/2023 號法律 醫學輔助生殖技術," 2023).

level of protection, the public authority may authorize such transfers, if the data processor can provide sufficient safeguards to privacy protection and the fundamental rights and freedoms of individuals.[67]

The Personal Data Protection Bureau (PDPB), which is responsible for implementing the *Personal Data Protection Act*, has not yet issued specific guidelines on handling data protection issues in research involving extensive healthcare data. This lack of specification may leave the research community with uncertainties when designing and conducting studies that involve the use of personal data and cross-border transfers of personal data.

IV. Recommendations

Based on the legal challenges identified, governments in the two regions should make efforts to improve the transparency of their respective legal regimes and administrative approval procedures and catalyze health data use for AI advancement in medical research and healthcare. It is worth noting that a timely removal of the data localization requirement and setting reasonable research exemptions for data use for scientific purposes in Mainland China could contribute to more promising AI-enabled ART research outputs. Moreover, legal reconciliation should also be made to mitigate the significant legal differences between these two regions, especially considering the meticulous and stringent data protection regulations introduced in Mainland China.

It is also worth noting that the governments of two regions have begun to address these legal challenges. For instance, the joint "Implementation Guidelines on the Standard Contract for Cross-boundary Flow of Personal Information within the Guangdong-Hong Kong-Macao Greater Bay Area (Mainland, Macao)" was established in September 2024. According to the Guidelines, data processors can transfer personal information between the two regions pursuant to the Standard Contracts signed.[68] However, the Guidelines only apply to personal information processors/ recipients registered (for organizations) or located (for individuals) in Macau or the nine Mainland cities within the Guangdong-Hong Kong-Macao Greater Bay Area in Guangdong Province. In addition, the personal information being transferred cannot be made available to organizations or individuals outside the Guangdong-Hong Kong-Macao Greater Bay Area.[69] Moreover, scientists engaged in cross-border research are still obliged to obey legal requirements covering special data categories, such as genetic data and medical records. Hence, it is imperative that additional mechanisms are put in place that will encourage and facilitate data flow for research purposes between the two regions. This could promote health data communication and pave the way for further AI-enabled ART research collaborations.

---

[67] Legislative Assembly of Macau SAR., "Personal Data Protection Act (Law No. 8/2005) 澳門特別行政區 第 8/2005 號法律 個人資料保護法," 2005, Article 20.2.

[68] Cyberspace Administration of China, Economic and Technological Development Bureau of Macau SAR., and Personal Data Protection Bureau of Macau SAR., "Implementation Guidelines on the Standard Contract for Cross-boundary Flow of Personal Information within the Guangdong-Hong Kong-Macao Greater Bay Area (Mainland, Macao) (粵港澳大灣區(內地、澳門)個人信息跨境流動標準合同實施指引)," 2024, Article 2.

[69] *Ibid.*, Articles 2 and 4.

Part 2: Macau SAR and Mainland China – 2

Ethics Governance in Medical AI Research and Development in Shenzhen, China

Dr Ji Ping
Vice Director, Clinical Research Institute,
Shenzhen Peking University-Hong Kong University of Science and Technology
Medical Center, China

This presentation delves into the landscape of health Artificial Intelligence (AI) in China, citing information from the 2023 Health AI Index Report from the National Institute of Health Data Science at Peking University.[70] The report reveals a significant surge in medical AI technology. Since 2017, China has been at the forefront of global health AI clinical trials, with lung cancer, breast cancer, and stroke emerging as the top therapeutic areas. Notably, 58% of approved AI medical devices are concentrated in radiology and medical imaging.

Governance must evolve to address the dynamic nature of AI. The presentation navigates the regulatory framework governing AI in China's healthcare sector. It encompasses:

1. Fundamental regulations concerning data and data users, including the Cybersecurity Law (网络安全法), Data Security Law (数据安全法), Personal Information Protection Law (个人信息保护法) and data protection laws (数据保护法)[71].

2. Medical AI devices can only be marketed and deployed after registration/ filing and classification, with the National Medical Products Administration (NMPA) (in particular, the Center for Medical Device Evaluation of NMPA) providing oversight and issuing regulations such as "Key Points of Deep Learning Aided Decision-Making Medical Device Software Review" (深度学习辅助决策医疗器械软件审评要点), "Guiding Principles for Classification and Definition of AI Medical Software Products" (人工智能医用软件产品分类界定指导原则) and "Guiding Principles for the Registration Review of AI Medical Devices" (人工智能医疗器械注册审查指导原则).

3. The government is actively promoting the development of the AI industry and has incorporated ethical governance into its strategic planning. A series of AI ethical governance documents have been issued, including the New Generation AI Governance Principles (新一代人工智能治理原则), New Generation AI Code of Ethics (新一代人工智能伦理规范), Position Paper of China on Strengthening Ethical Governance of AI (中国关于加强人工智能伦理治理的立场文件), and Provisional Measures for the Administration of Generative AI Services (生成式人工智能服务管理暂行办法).

---

[70] Zhan, Qimin and Dong Erdan. "Health Artificial Intelligence Index Report 2023 (健康医疗人工智能指数报告 2023)" [M]. 1st ed. Beijing: Science Press, 2024.
[71] The legal regime of data protection of Mainland China in general.

4. The “Opinions on Strengthening Ethical Governance of Science and Technology” (关于加强科技伦理治理的意见) emphasized ethical governance in life sciences, medical research, and AI fields. The latest significant national regulations issued by, among others, the Ministry of Science and Technology and the National Health Commission included the “Provisional Measures for Ethics Review of Science and Technology” (科技伦理审查办法试行) and the “Ethics Review Measures of Life Science and Medical Research Involving Human Subjects” (涉及人的生命科学和医学研究伦理审查办法).

Under the national ethical governance structure, AI governance falls under the Ministry of Industry and Information Technology, which has established an Ethical Expert Committee. The China Academy of Information and Communications Technology supports this committee in providing decision-making advisory support, conducting expert reviews, developing regulations, standards, and guidance, providing training, and evaluating institutional ethics review committees. Currently, AI regulatory bodies are established at the national, provincial, and municipal levels, but the regulatory responsibilities are fragmented, involving multiple independent government departments in charge of scientific and technological innovation, health, industry, and information technology.

As China's inaugural “Special Economic Zone” and a “City of Innovation”, Shenzhen is a vanguard in promoting “AI + medical health.” Extensive development of AI-related healthcare products is underway, involving government, industry, research institutes, hospitals, and universities. However, implementing ethical principles in practice remains a formidable challenge. Our objectives are to assess the current status, ethical challenges, barriers, and needs, and to inform policymakers on ethical governance. Stakeholders have identified favorable drivers for medical AI R&D in Shenzhen, including increased governmental funding, supportive policies, and a widespread recognition of AI's value and demand in healthcare. AI's empowerment of the medical field has brought benefits to hospital management, clinical diagnosis and treatment, nursing and health services, and medical research. However, various ethical challenges associated with emerging AI-based technologies have been identified:

1) Local government authorities are most concerned with regulatory aspects, robust regional ethical governance, and the hierarchical management of ethical risks. At the medical institution level, AI regulation often involves the medical affairs department, equipment management department, information department, ethics committee, and related departments. Interviewees raised questions such as “How to conduct a rationality and applicability assessment for access, and implement an evaluation mechanism and process for AI medical products and services?”, “How to evaluate the application effects in terms of utilization rate and impact on business processes, hospitals, and healthcare workers?”

2) Designers and researchers face compliance issues, ethical product design and application, and the acquisition of high-quality data. Interviewees reported that “obtaining and sharing high-quality data is extremely difficult. Companies are trying all kinds of methods and spending a lot of costs, but the amount of data may not be large enough, and the quality of the data may not be guaranteed.”

3) Doctors and hospital users may over-rely on or distrust AI technology, with concerns about reduced human interaction, affecting doctor-patient relationships, unreliable diagnostic results, and unclear medical liability. Feedback from the radiology interviewee indicates that AI has become an indispensable tool in daily work, which might weaken doctors’ own diagnostic abilities and potentially hinder their progress in knowledge acquisition and professional skill development, which is detrimental to the training and growth of young doctors.

AI must enhance health system capacity and advance patient interests. It must be rigorously assessed for safety and efficacy in clinical settings. Proposed strategies include developing specific work guidelines, a risk-based grading supervision model, and a joint governance mechanism for local AI government authorities. During the development phase, a thorough review for safety, effectiveness, and ethics should be conducted, along with a comprehensive risk-benefit and cost-effectiveness evaluation. Medical institutions must establish robust medical information systems, equip staff with professional training, and integrate AI applications into clinical decision-making processes. Strengthening healthcare institutional capacity is crucial through improved management, regulatory compliance, proper data governance, and fostering interdisciplinary collaboration. Regulation, practice standards, and guidance should be developed and updated based on evidence, with a focus on real-world AI technology implementation.

Collaboration between regulatory authorities, technology developers, and users is imperative for the development and application of AI technology. Shenzhen Health Commission issued the "Notice on the Safety Management of 'AI+Healthcare' Application" ("AI+医疗健康"应用安全管理工作的通知) in 2024, which clearly requires that designers, manufacturers, and operators should establish a mechanism for sharing responsibilities. With effective governance policies and a well-established ecosystem, Shenzhen can harness its potential as a hub for healthcare AI innovation.

Part 3: European Union and the Council of Europe – 1

Medical Device Compliance under the EU AI Act and the EU's Medical Device Regulation

Prof Timo Minssen
LL.D., LL.Lic., LL.M., M.I.C.L., Dipl. Jur.,
Professor, Founder & Director, Centre for Advanced Studies in Bioscience Innovation Law (CeBIL), Faculty of Law,
UNESCO Co-Chair in the Right to Science,
University of Copenhagen, Denmark;
Research Affiliate, Centre for Law, Medicine and Life Sciences (LML),
University of Cambridge, United Kingdom;
Inter-CeBIL Affiliate, Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics,
Harvard University, Cambridge, MA, United States

The European Union's Artificial Intelligence Act (EU AI Act) represents a landmark effort to regulate Artificial Intelligence (AI) technologies.[72] Its overarching goals emphasize enhancing safety, transparency, and accountability for AI systems, aiming to establish a global benchmark for AI regulation. While the EU AI Act seeks to promote ethical considerations and responsible deployment of AI technologies, it also seeks to create a regulatory environment that promotes innovation. However, achieving this dual objective poses significant challenges and risks, particularly in balancing innovation with regulatory compliance. This is also particularly relevant in the context of regulated digital healthcare products, including medical devices that need to comply with the Medical Device Regulation (MDR)[73] and the Regulation on In Vitro Diagnostic Medical Devices (IVDR) both of the European Union (EU).[74]

This presentation was based on the recent paper "Navigating the EU AI Act: Implications for Regulated Digital Medical Products" by Mateo Aboy, Timo Minssen and Effy Vayena, published in *npj Digital Medicine*.[75] Our paper discussed the potential implications of the EU AI Act on the development of AI driven medical products and its interplay with the EU MDR. Special emphasis was placed on the challenges it poses for stakeholders, and the ongoing need for regulatory adaptability in the face of rapid technological advancements. The following sections provide a brief overview on the main topics and findings of this presentation:

1. The Dual Objectives of the EU AI Act

At its core, the EU AI Act aspires to encourage the development of advanced AI and machine learning (ML) systems, as well as general purpose AI models, including those that might – or might not - qualify as medical devices, while ensuring compliance with stringent safety and efficacy standards. However, the EU AI Act introduces a layer of complexity, particularly for developers and providers of AI/ML medical devices, who must navigate the regulatory requirements of both the EU MDR and

[72] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), PE/24/2024/REV/1, OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj.

[73] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on Medical Devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002, and Regulation (EC) No 1223/2009, and Repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017: 1–175.

[74] Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU, OJ L 117, 5.5.2017: 176–332.

[75] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7, no. 237 (2024). https://doi.org/10.1038/s41746-024-01232-3.

the EU AI Act. While there are areas of overlap between these regulations—such as a risk category-based assessment, quality management systems (QMS), and post-marketing surveillance — various definitions and specific distinctions can lead to confusion and increased regulatory burdens, particularly for small and medium-sized enterprises (SMEs).[76]

SMEs often face challenges in allocating limited resources to meet regulatory demands, which can detract from their core focus on engineering, quality assurance, and product development.[77] Medical device companies, and in particular SMEs, are struggling with the additional documentation requirements, increased costs, and a lack of clarity in regulatory expectations. The EU recognized this challenge at a time of increasing international competition and has recently initiated attempts to streamline and simplify regulations and compliance requirements. Similar considerations have also resulted in the recent call by the EU Parliament to reform the EU MDR[78] and the recent withdrawals of forthcoming EU legislation on AI liability.[79] These events have demonstrated, that it is a very delicate task to balance the risks of over- and under-regulation to protect citizens' rights, EU values, and the competitiveness of the EU at the same time.[80] These challenges with the EU MDR are exacerbated by the EU AI Act, which adds further regulatory obligations, potentially stifling innovation and leading to the withdrawal of products from the market. The EU's current status as a leader in medical AI patents could be threatened if the EU AI Act creates barriers that disproportionately impact SMEs, which form an important part of the European economy.[81]

2. Stakeholder Collaboration and Regulatory Capacity

The successful implementation of the EU AI Act hinges on the collaboration of a diverse array of stakeholders, including policymakers, regulators, notified bodies, AI providers, and the public sector. The potential for bottlenecks arises if any stakeholder fails to meet the EU AI Act's requirements. For instance, the EU's implementation of the MDR has been hampered by a shortage of notified bodies, leading to compliance delays and legal uncertainties. As most AI/ML-enabled medical devices are classified as high-risk, the increased demand for notified body assessments could overwhelm existing capacities.

Additionally, the establishment of simplified technical documentation for SMEs and the development of AI codes of practice by the AI Office are still nascent[82] or pending, raising concerns about the overall readiness of the regulatory framework to support the EU AI Act's objectives. In this environment, improving the coordination between the various stakeholders and increasing the capacity of regulators, decision makers and users must be a major objective.

---

[76] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7, no. 237 (2024). https://doi.org/10.1038/s41746-024-01232-3.
[77] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7, no. 237 (2024). https://doi.org/10.1038/s41746-024-01232-3.
[78] European Parliament resolution on the urgent need to revise the Medical Devices Regulation (2024/2849(RSP)), available at: https://www.europarl.europa.eu/doceo/document/RC-10-2024-0123_EN.html.
[79] PYMNTS, "European Commission Withdraws Draft Rules on Technology Patents, AI Liability and Consumer Privacy" (12 February 2025). https://www.pymnts.com/news/regulation/2025/european-commission-withdraws-draft-rules-on-technology-patents-ai-liability-and-consumer-privacy/.
[80] See e.g. Tallberg, Jonas, Magnus Lundgren, and Johannes Geith, "AI Regulation in the European Union: Examining Non-State Actor Preferences", *Business and Politics* 26, no. 2 (2024): 218–239, https://doi.org/10.1017/bap.2023.36.
[81] Tallberg, Jonas, Magnus Lundgren, and Johannes Geith, "AI Regulation in the European Union: Examining Non-State Actor Preferences", *Business and Politics* 26, no. 2 (2024): 218–239, https://doi.org/10.1017/bap.2023.36.
[82] See e.g. European Commission, "Commission Guidelines on Prohibited Artificial Intelligence (AI) Practices Established by Regulation (EU) 2024/1689 (AI Act)". C(2025) 884 final, Brussels (4 February 2025), https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act; see also: European Commission, "Guidelines on the Definition of an Artificial Intelligence System Established by the AI Act" (2025), https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application.

3. Addressing Generative AI and the "Intended Use" Challenge

The emergence of large multimodal models poses additional regulatory challenges, particularly regarding their intended use. Many AI models, including general-purpose AI, can be repurposed for multiple applications, complicating their classification under existing regulations. For example, a medical AI device designed for image analysis could be enhanced by integrating a large language model (LLM) to improve diagnostic reasoning and output capabilities.

The EU AI Act includes specific provisions on such General Purpose AI (GPAI) applications, but the interplay between the EU MDR and the EU AI Act becomes increasingly complex in such scenarios, highlighting the need for clarity in regulatory definitions and classifications.[83] Furthermore, the efficacy of LLMs in health and life science innovation may to some extent depend on legal compliance, including copyrights, which is why it has been argued that the future for LLMs in medicine must be based on transparent and controllable open-source models.[84]

4. The Need for Adaptive Regulation

The rapid pace of AI innovation necessitates a flexible and adaptive regulatory framework. The initial version of the EU AI Act did not account for the rise of generative AI, illustrating the challenges of keeping legislation aligned with technological advancements. The last-minute inclusion of provisions addressing generative AI underscores the difficulties of creating static regulations in a dynamic environment.[85]

Regulatory frameworks must be able to accommodate emerging technologies, methodologies, and ethical considerations while remaining responsive to societal expectations. The EU AI Act's comprehensive risk-based approach aims to address these challenges, but its effectiveness will depend on ongoing adjustments and improvements based on real-world experiences and outcomes.[86] Regulatory sandboxes, as provided for in Article 57 of the EU AI Act, can be an important element in this endeavor, but it will be important to coordinate their implementation throughout Europe.

5. Contrasting Approaches: The United Kingdom as a Case Study

The approach of the United Kingdom (UK) to AI regulation offers an interesting contrast to the EU's comprehensive legislative framework. Post-Brexit, the UK has sought to position itself as a leader in AI by adopting a flexible, pro-innovation regulatory environment. The UK government emphasizes

---

[83] European Commission, "Commission Guidelines on Prohibited Artificial Intelligence (AI) Practices Established by Regulation (EU) 2024/1689 (AI Act)". C(2025) 884 final, Brussels (4 February 2025), https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act; see also: European Commission, "Guidelines on the Definition of an AI System Established by the AI Act" (2025), https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application; see also Minssen, Timo, Effy Vayena and I. G. Cohen, "The Challenges for Regulating Medical Use of ChatGPT and Other Large Language Models", *JAMA* 330, no. 4 (2023): 315-316. https://doi.org/10.1001/jama.2023.9651.

[84] Riedemann, Lars, Maxime Labonne and Stephen Gilbert, "The Path Forward for Large Language Models in Medicine Is Open", *npj Digital Medicine* 7 (2024): 339. https://doi.org/10.1038/s41746-024-01344-w.

[85] Mann, Sebastian Porsdam, I. Glenn Cohen, and Timo Minssen, "The EU AI Act: Implications for U.S. Health Care", *NEJM AI* 1, no. 11 (2024). http://doi.org/10.1056/AIp2400449. See also: Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7 (2024): 237. https://doi.org/10.1038/s41746-024-01232-3.

[86] Mann, Sebastian Porsdam, I. Glenn Cohen, and Timo Minssen, "The EU AI Act: Implications for U.S. Health Care", *NEJM AI* 1, no. 11 (2024). http://doi.org/10.1056/AIp2400449. See also: Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7 (2024): 237. https://doi.org/10.1038/s41746-024-01232-3.

principles-based regulation, focusing on safety, transparency, and accountability while allowing for sector-specific adaptations by existing regulators.

The UK's non-statutory guidance encourages regulators to promote innovation and competition, fostering an environment that is responsive to the rapidly evolving nature of AI technologies. This pragmatic approach aims to mitigate the risk of stifling innovation through rigid legislative requirements, enabling the UK to adapt more rapidly to emerging AI developments.[87]

Similar alternative approaches to AI and medical device regulation are being developed in other countries, such as the United States, China, Canada, Switzerland and South Korea and it is going to be interesting and important to analyze, compare, and evaluate these approaches.

6. Conclusion

The EU AI Act represents a significant effort to regulate AI technologies in healthcare, balancing the need for safety and efficacy with the imperative to foster innovation. However, the potential challenges posed by overlapping regulatory requirements, stakeholder collaboration, and the rapid pace of technological change cannot be overlooked. The complexity of the rapidly evolving EU regulatory landscape, such as the recently enacted European Health Data Space Regulation[88] and further legislation, as well as the aforementioned recent attempts by the EU to streamline regulation and make compliance easier, also means that the EU AI Act, its interpretation, and its implementation will have to evolve. Continuous monitoring and research are therefore essential to assess the EU AI Act's impact on innovation, particularly concerning SMEs and the overall competitive landscape in digital healthcare.

To ensure that the EU AI Act achieves its intended goals, a proactive and evidence-based approach to regulation is necessary. Ongoing stakeholder engagement, adaptive legislative frameworks, and a willingness to learn from real-world applications will be crucial in navigating the complexities of AI regulation.[89] Ultimately, the success of the EU AI Act will depend on its ability to evolve alongside technological advancements and regulatory developments while maintaining public trust and compliance in the healthcare sector.[90]

---

[87] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7 (2024): 237. https://doi.org/10.1038/s41746-024-01232-3, citing: Department for Science, Innovation & Technology of the United Kingdom, "Implementing the UK's AI Regulatory Principles: Initial Guidance for Regulators", (2024) https://assets.publishing.service.gov.uk/media/65c0b6bd63a23d0013c821a0/implementing_the_uk_ai_regulatory_principles_guidance_for_regulators.pdf; Secretary of State for Science, Innovation and Technology of the United Kingdom, "A Pro-innovation Approach to AI Regulation", (2023) https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper; UK Medicines & Healthcare products Regulatory Agency (MHRA), "Guidance-Software and Artificial Intelligence (AI) as a Medical Device", (February 2025). https://www.gov.uk/government/publications/software-and-artificial-intelligence-ai-as-a-medical-device/software-and-artificial-intelligence-ai-as-a-medical-device.

[88] Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, PE/76/2024/REV/1, OJ L, 2025/327, 5.3.2025, ELI: http://data.europa.eu/eli/reg/2025/327/oj .

[89] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7 (2024): 237. https://doi.org/10.1038/s41746-024-01232-3.

[90] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7 (2024): 237. https://doi.org/10.1038/s41746-024-01232-3.

Part 3: European Union and the Council of Europe – 2

The European Health Data Space: Realistic Ambitions for Health Research and Innovation?

Dr Colin Mitchell
Head of Humanities, PHG Foundation, University of Cambridge, United Kingdom

The European Health Data Space (EHDS) presents an ambitious step towards a more integrated and streamlined approach to data access for health research and innovation across the European Union (EU). It has the potential to significantly accelerate scientific progress, boost collaboration, and facilitate knowledge transfer between different scientific disciplines.[91] However, with great ambition comes considerable challenges which EU bodies and member states will need to overcome to achieve this potential.

Scope and ambition of the EHDS

What is the EHDS?

The EHDS is one of nine common data spaces envisaged by the 2020 European Strategy for Data,[92] focused on EU priority areas such as health, agriculture and manufacturing. The aim is to allow data to be made available and exchanged in a trustworthy and secure manner across the EU. This requires the development of common data infrastructure and governance frameworks.

The European Health Data Space Regulation, which entered into force on 25 March 2025,[93] sets out the infrastructure, governance and standards that will apply in the context of electronic health data. This applies to two uses, primary use and secondary use. Primary use means the processing of electronic health data for the provision of healthcare, in order to assess, maintain or restore the state of health of the natural person to whom those data relate, including the prescription, dispensation and provision of medicinal products and medical devices, as well as for relevant social, administrative or reimbursement services. Secondary use means the processing of electronic health data for the purposes set out in Chapter IV of the European Health Data Space Regulation (such as research, innovation, policy-making, and the public interest in the areas of public or occupational health as more particularly described in Chapter IV), other than the initial purposes for which they were collected or produced.

The focus of this contribution is on the secondary use element of the EHDS, in particular, its potential impact on health research and Artificial Intelligence (AI) development.

How does the EHDS seek to promote research and innovation?

The essential elements of the EHDS for research and innovation include the establishment of a cross-border infrastructure for secondary use of electronic health data (HealthData@EU), and a common mechanism for access to electronic health data for secondary use. The latter involves:

- Application to a health data access body (HDAB) for access to electronic health data

[91] European Commission, "European Health Data Space Regulation (EHDS)." https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en.
[92] European Commission, "Common European Data Spaces – Shaping Europe's Digital Future", https://digital-strategy.ec.europa.eu/en/policies/data-spaces.
[93] Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847.

- If approved, mandatory disclosure of personal and non-personal electronic health data by *data holders*, for processing by *data users* within a secure processing environment
- Fees collected by the HDAB or trusted health data holder proportionate to the cost of making the data available
- All subject to potential individual opt-out and specific member state safeguards and additional restrictions

The EHDS creates a whole new scheme of actors and bodies with different entitlements and responsibilities. Notably, health data access bodies are required to be established within each member state and bear the greatest responsibility for ensuring the smooth operation of the EHDS in their jurisdiction.

What data should be made available and by whom?

The EHDS is based on a very broad understanding of "electronic health data that can be processed for secondary use". Article 51 of the Regulation on the European Health Data Space sets out a range of categories including – amongst others – electronic health data from electronic health records (EHRs), data on socio-economic and environmental determinants of health, human genomic data, clinical trials data, health data from biobanks and data from medical devices and wellness applications. Importantly, this encompasses a range of data initially collected for other purposes and there is also provision for Member States to add to this list.

There is a similarly broad interpretation of bodies that may be a source of data for secondary use in the EHDS – "health data holders" (defined in Article 2(2)(t)). This does not just include any natural or legal person, public authority, agency or other body in the healthcare or the care sectors but also any natural or legal person developing products or services intended for the health, healthcare or care sectors, developing or manufacturing wellness applications, performing research in relation to the healthcare or care sectors or acting as a mortality registry, as well as any Union institution, body, office or agency, that has either:

(i) "the right or obligation … to process personal electronic health data for the provision of healthcare or care or for the purposes of public health, reimbursement, research, innovation, policymaking, official statistics or patient safety or for regulatory purposes"; or

(ii) "the ability to make available non-personal electronic health data through the control of the technical design of a product and related services …"

Overall, a wide range of bodies *must* make a diversity of electronic health-related data available through the EHDS if requested for permitted secondary purposes.

Which secondary purposes are permitted?

The EHDS enables secondary use of electronic health data for health or care research purposes. Crucially, in the context of medical AI, Article 53(1)(e), this explicitly incorporates: "(i) development and innovation activities for products or services; (ii) training, testing and evaluating of algorithms, including in medical devices, *in-vitro* diagnostic medical devices, AI systems and digital health applications".

While existing modes of data access are not ruled out by the new Regulation, access to the EHDS is intended to be more cost-effective, streamlined, and powerful – unlocking swathes of data that would otherwise not be readily available to some users.

Challenges

However, there are several significant challenges for those seeking to implement the EHDS and those established in "third countries" who may wish to interact with it.

Mandatory data disclosure v opt out

A topic of great debate during the development of the EHDS regulation was the proposal for a system of mandatory data disclosure for approved secondary purposes. The desire behind this is to ensure that the data within the EHDS is as comprehensive and complete as possible to power scientific research and innovation. However, scholars and patient representatives argued that this approach was unethical, would potentially counter existing patient or data subject choices, and would unjustifiably infringe individual autonomy.[94] It was also argued that this could ultimately impact public trust and support for the EHDS endeavour.

The final agreed approach provides individuals with a right to opt out of the use of data for secondary purposes at any time (Article 71). However, member states may establish overrides for situations in which the processing of the data is necessary for scientific research "for important reasons of public interest" or any of the other specified purposes, where the health data access application or health data request is submitted by a public sector body or a Union institution, body, office or agency with a mandate to carry out tasks in the area of public health, or by another entity entrusted with carrying out public tasks in the area of public health, or acting on behalf of or commissioned by a public authority and the other conditions are met (Article 71(4)). While a high bar is set for such overrides, it may be that this will introduce variation in the availability of data across the EU.

As well as the potential impact on completeness of the data contained in the EHDS, the challenge of this final position is that providing an opt-out by no means guarantees public confidence and support, in particular if the initiative is poorly publicised. For example, this has been the repeated conclusion of population level health data initiatives in the UK.[95]

In addition, a range of commercial data holders also objected to the mandatory nature of the EHDS regime, arguing that this may infringe their intellectual property rights and require disclosure of commercially sensitive data.[96] Despite lobbying there has been no opt-out provided for data holders but the HDABs shall refuse access to electronic health data for secondary use where the grant of such access entails a serious risk of infringing intellectual property rights, trade secrets or the regulatory data protection right specified which cannot be addressed in a satisfactory manner. The challenge with this is the level of expertise and capacity that the HDAB will require to properly consider such requests and the potential that commercial organisations may challenge their decisions, including through litigation. Moreover, the success of the EHDS requires the support of data holders as much as it rests on the continued confidence of data subjects.

Capacity and resources

To achieve the scale of the EHDS requires enormous resources – in particular to develop capacity and digital infrastructure across the EU despite highly variable existing states of digital

[94] Staunton, Ciara, Mahsa Shabani, Deborah Mascalzoni, Signe Mežinska, and Santa Slokenberga. "Ethical and Social Reflections on the Proposed European Health Data Space." *European Journal of Human Genetics* 32, no. 1 (2024): 498-505. https://doi.org/10.1038/s41431-024-01543-9.

[95] Redhead, Caroline A. B., Catherine Bowden, John Ainsworth, Nigel Burns, James Cunningham, Søren Holm, and Sarah Devaney. "Unlocking the Promise of UK Health Data: Considering the Case for a Charitable GP Data Trust." *Medical Law Review* 33, no.1, January 4, 2025. https://doi.org/10.1093/medlaw/fwae043.

[96] Redrup Hill, Elizabeth. "The European Health Data Space." *PHG Foundation*, October 2023. https://www.phgfoundation.org/wp-content/uploads/2023/10/The-European-Health-Data-Space-briefing.pdf.

readiness. The Commission will provide over €810 million to support the EHDS and further funding sources will provide additional millions, but it remains to be seen if that level of funding will prove adequate to create a seamless infrastructure and governance system across the EU.

Frictions with wider legal and ethical framework

Another challenge which will need to be addressed to achieve the EHDS' aims is overcoming friction with wider regulatory and ethical frameworks relating to data and research. For example, the General Data Protection Regulation (GDPR) still applies in full to processing of personal data within the EHDS and the EHDS – to an extent – adds to the complexity of compliance by enabling further conditions by Member States for international access to data and processing of sensitive data (under Articles 90 & 51(4) EHDS Regulation respectively).

In addition, the EHDS does nothing to harmonise and streamline research ethics requirements which may still apply by virtue of variable national law. Ultimately a significant level of variation and complexity will continue to apply across the EU.

Challenges for international collaboration

Finally and perhaps most crucially for those outside the EU/ European Economic Area (EEA), what scope is there for access to the EHDS by "third country" innovators and researchers? The EHDS position is that such access is feasible but only on the basis of reciprocity. Article 91 of the EHDS Regulation provides that health data access applications and health data requests submitted by a health data applicant established in a third country shall be considered eligible by health data access bodies and the Union health data access service if the third country concerned (not simply the requesting institution) (i) is an authorised participant on the basis of having a national contact point for secondary use covered by an implementing act referred to in Article 75(5) or (ii) allows Union health data applicants access to electronic health data in that third country under conditions that are not more restrictive than those provided for in this Regulation, and therefore such access is covered by an implementing act referred to in Article 91(2). Recital (94) of the EHDS Regulation states that "[m]aking electronic health data available to a third country should be allowed to take place only where the Commission has established … that the third country concerned allows access to electronic health data originating from that third country by Union entities under the same conditions and with the same safeguards as would be the case if they were accessing electronic health data within the Union". Moreover, the GDPR will continue to apply its very high standards to requests for access to personal data within the EHDS from third countries. Taken together, it is clear that international collaboration and participation in the EHDS will not be easily achieved.

Conclusions

The EHDS represents a hugely ambitious step toward streamlining and unlocking a diversity of electronic health data (defined broadly) for research, innovation and AI development. However, the capacity and infrastructure development requirements are very significant and maintaining wider patient, public, member state and industry support will also be key. EHDS requirements cause potential challenges for all these groups. For those outside the EU/EEA- and as with the GDPR and AI Act, the EHDS *may* inspire global approaches, particularly given access is contingent on reciprocal arrangements. However, whether or not this proves too burdensome for other countries remains to be seen and there is a significant prospect of global divergence – as opposed to convergence – in this field.

Part 3: European Union and the Council of Europe – 3

The Council of Europe's AI Convention and its Potential Impact on Healthcare

Prof dr Tom Goffin
Associate Professor of Health Law, Metamedica
Ghent University, Belgium

1. Introduction

In the context of Europe's regulatory frameworks concerning Artificial Intelligence (AI) in healthcare, the primary subject of discussion is Regulation (EU) 2024/1689 of 13 June 2024, the so-called AI Act (EU AI Act).[97] This regulatory framework occupies a central role in the broader discourse on AI governance. The EU AI Act is a pivotal document within the legal framework of AI in healthcare in Europe, but it is not the sole international document of significance in this regard.

The legislative framework of Europe is not solely comprised of the legislation of the European Union (EU); the legislation of the Council of Europe is also a contributing element. The Council of Europe, an international organization, is dedicated to the promotion of human rights, democracy, and the rule of law in Europe. Established in 1949 as a response to the Second World War, it stands as Europe's oldest intergovernmental organization, comprising 46 member states from across the continent. In contrast to the EU, the Council of Europe lacks the authority to enact legally binding laws. However, it has played a significant role in the development of international treaties, including the 1953 Convention for the Protection of Human Rights and Fundamental Freedoms, also known as the European Convention on Human Rights (ECHR). Provisions from the convention are incorporated into domestic law in the Contracting Member States. Within the Council of Europe, the European Court of Human Rights (ECtHR) has the authority to adjudicate alleged violations of the European Convention on Human Rights by member states. In biomedicine and healthcare, the Council of Europe has undertaken numerous initiatives over the years, most notably the European Convention on Human Rights and Biomedicine, also known as the Patients' Rights Convention, in addition to several additional protocols.[98]

On 17 May 2024, the Committee of Ministers for the Council of Europe adopted the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, also known as the AI Convention.[99] This new framework aligns with the Council of Europe's established standards, promoting innovation and comprising a binding legal instrument of a transversal character,

[97] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), OJ 12 July 2024.
[98] Council of Europe, Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (ETS No. 164), entry into force 1 December 1999.
[99] Council of Europe, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, open for signature 5 September 2024, https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.

encompassing general common principles.[100] In contrast to the provisions of the EU AI Act, it does not explicitly regulate the economic and market aspects of AI systems. The Council of Europe lacks the competence to regulate these economic and market aspects. The document, when considered in its entirety, establishes a common legal framework at the global level. It aims to ensure that the activities within the lifecycle of AI systems by public and possibly private actors comply with existing international and domestic legal obligations, standards, and commitments in the spheres of human rights, democracy, and the rule of law.

Within the Convention, AI is defined as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment." (Article 2). The text of the Convention establishes a range of obligations for the Contracting Parties related to the activities throughout the lifecycle of AI systems. These activities include the following: (1) planning and design, (2) data collection and processing, (3) development of AI systems, including the construction of models and/or refinement of existing models for specific tasks, (4) testing, verification, and validation, (5) supply/availability for use, (6) deployment, (7) operation and monitoring, and (8) retirement. Applying these obligations throughout the lifecycle ensures the Convention can handle current and future risks, especially given the rapid and often unpredictable nature of technology. The Convention's horizontal application covers public and possibly private sectors (depending on the implementation by the Contracting Parties, with limited exceptions for national security, research and development (unless they could potentially interfere with human rights, democracy, and the rule of law), and national defence. (Article 3). In comparison to the scope of the EU AI Act, the AI Convention's field of application is broader.

Like the Biomedicine Convention, the Council of Europe has chosen the form of a Framework Convention for the AI Convention. Whilst the Framework Convention is a legally binding instrument under international law, the word "Framework" highlights the scope for Contracting Parties to translate the AI Convention's provisions to their specific country situation through national legislation and appropriate governmental policies.[101] The AI Convention ensures the application of obligations related to human rights, democracy, and the rule of law—including the Charter of Fundamental Rights of the EU within EU Member States—to activities related to AI systems. It aligns with the human rights protection systems and mechanisms of each Contracting Party and their international law obligations. The AI Convention does not establish new human rights or obligations or diminish existing protections. It establishes various legally binding obligations in Chapters II to VI to facilitate the implementation of each Party's applicable human rights obligations in the context of new challenges posed by AI. The AI Convention is unmistakably significant despite its lack of novel rights

[100] Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 2024, paragraph 2, https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.

[101] A definition of what the Council of Europe means by a "framework convention", was provided in paragraph 11 of the Explanatory Report to the Council of Europe's Framework Convention for the Protection of National Minorities (ETS No. 157), entry into force 1 February 1998.

or obligations. It specifies how AI should be used in line with the Council of Europe's human rights standards.[102]

2. State's obligations

According to Article 1 of the Convention, the Contracting Parties must implement appropriate legislative, administrative, or other measures to implement the provisions outlined in the Convention. These measures should be graduated and differentiated based on the severity and likelihood of adverse impacts on human rights, democracy, and the rule of law throughout the lifecycle of AI systems. This may include specific or horizontal measures that apply irrespective of the type of technology used. As the Explanatory Report to the AI Convention indicates, the Convention's implementation is an obligation of result, not of means. In this respect, the principle of subsidiarity is essential, as it gives the Parties the primary responsibility to ensure respect for human rights and to provide redress for human rights violations.[103]

The Convention stipulates general obligations – protection of human rights and integrity of democratic processes and respect for the rule of law – and general principles related to activities within the lifecycle of AI systems – respect for human dignity and individual autonomy, transparency and oversight, accountability and responsibility, equality and non-discrimination, privacy and personal data protection, reliability and safe innovation – for actors applying AI systems (Chapter II and III). It thereby formalises fundamental AI principles to underpin concrete individual rights.[104]

In alignment with the principles outlined in the EU AI Act, the Convention proposes a risk and impact management framework aiming to ensure that AI systems do not infringe upon human rights. Article 16 stipulates the obligation of the Contracting Parties to adopt measures for the identification, assessment, prevention, and mitigation of risks posed by AI systems, taking into account their actual and potential impacts on human rights, democracy, and the rule of law. These measures are also expected to ensure that any adverse effects of AI systems are adequately addressed and documented.

The Convention also stipulates some specific obligations for Member States implementing the Convention, such as ensuring non-discrimination (Article 17), providing for public consultation (Article 19), and investing in digital literacy and skills (Article 20). Finally, the Convention requires States to report to the Conference of Parties with details of activities undertaken to give effect to the implementation of the Convention (Article 24), the exchange of relevant and useful information between Contracting Parties (Article 25), and to establish effective oversight mechanisms (Article 26).

The Convention opened for signatures on 5 September 2024, at the Conference of Ministers of Justice in Vilnius, Lithuania. It will enter into force 3 months after it is agreed to by 5 signatories, including

[102] Van Kolfschooten, Hannah and Carmel Shachar, "The Council of Europe's AI Convention (2023–2024): Promises and Pitfalls for Health Protection", *Health Policy* 138 (2023): 2-5. https://doi.org/10.1016/j.healthpol.2023.104935.
[103] Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, September 2024. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.
[104] Van Kolfschooten, Hannah and Carmel Shachar, "The Council of Europe's AI Convention (2023–2024): Promises and Pitfalls for Health Protection", *Health Policy* 138 (2023): 2-5. https://doi.org/10.1016/j.healthpol.2023.104935.

3 Council of Europe member states (Article 30). The EU will implement the Convention through the EU AI Act.[105] It will be implemented alongside other relevant EU laws when needed. Since the EU AI Act and the AI Convention apply to slightly different areas, it is important to assess the adequacy of the EU's implementation of the Convention in the EU AI Act.

3. Relevance for Healthcare

Although the impact of the Council of Europe's work is clear for past initiatives, such as the Biomedicine Convention and the relevant case law of the ECtHR,[106] it remains unclear what the exact impact of the AI Convention on healthcare once entered into force will be. Many questions remain unclear, such as: Will individual countries sign and ratify the convention, or is the signature of the EU sufficient, specifically in protecting human rights, democracy and the rule of law within the healthcare domain? How will Contracting Parties implement the Framework Convention in the field of health? What with Health AI that does not fall under the EU AI Act but does fall under the AI Convention (lifecycle)?[107] Will the ECtHR be competent to rule on health AI applications, for example, not respecting fundamental patient rights?[108] Will there be a link with the Biomedicine Convention and its additional protocols?

Although the AI Convention has the potential to improve health and patients' rights globally, whether this potential will be met will depend on how the AI Convention is implemented by its Contracting Parties.

[105] European Commission, "Commission Sings Council of Europe Framework Convention on Artificial Intelligence", Press Release, 5 September 2024, https://digital-strategy.ec.europa.eu/en/news/commission-signs-council-europe-framework-convention-artificial-intelligence#:~:text=The%20Commission%20has%20signed%20the,on%20behalf%20of%20the%20EU.&text=The%20Convention%20is%20the%20first,AI%20regulation%20in%20the%20world.
[106] European Court of Human Rights, "Factsheet. Health", September 2024. https://www.echr.coe.int/documents/d/echr/fs_health_eng.
[107] Palmieri, Sofia and Tom Goffin, "A Blanket that Leaves the Feet Cold: Exploring the AI Act Safety Framework for Medical AI", *European Journal of Health Law* 30, no. 4: 406-427, http://doi.org/10.1163/15718093-bja10104.
[108] Van Kolfschooten, Hannah and Carmel Shachar, "The Council of Europe's AI Convention (2023–2024): Promises and Pitfalls for Health Protection", *Health Policy* 138 (2023): 3-5. https://doi.org/10.1016/j.healthpol.2023.104935.

Part 4: UK

The UK Approach to Regulating AI (in Healthcare): Innovation First, Safety Second?

Prof Oliver Quick
Professor of Health Law and Policy, University of Bristol, United Kingdom

In February 2024, the UK Government confirmed its commitment to a "pro innovation" approach to regulating Artificial Intelligence (AI).[109] This non-statutory sector-specific model adopts a "wait and see" position towards managing the risks associated with AI. This light touch approach diverges from the extensive risk-based legislative framework produced by the European Union (EU).[110] Instead of creating an AI-specific super regulator or introducing new legal rules, the UK system is based on soft compliance with the following cross sectoral principles: safety, security and robustness; appropriate transparency and explainability; fairness; accountability and governance, and contestability and redress. This note briefly considers some key questions about the relationship between innovative technologies and healthcare regulation in the UK context: how will this principles-based approach to regulation be delivered? Is this agile regulation or active de-regulation which unduly prioritises digital industrial policy goals? Can current mechanisms ensure sufficient safety and adequate redress for AI related healthcare harm?

Regulation can be defined narrowly (with reference to law and regulators) or more broadly as "the sustained and focused attempt to alter the behaviour of others according to defined standards and purposes with the intention of producing a broadly identified outcome or outcomes, which may involve mechanisms of standard-setting, information-gathering and behaviour modification."[111] This broader sense is more relevant to the challenge of regulating something as vast, diverse and dynamic as AI. Whilst claims to innovation are often overplayed,[112] there is no doubting the disruptive implications of innovation in altering attitudes towards harm and preventability, requiring professional (and patient) education and training and raising questions about how law and regulation can evolve alongside technology.[113] There are many different regulatory approaches – "command & control", "co-regulation", "self-regulation" and a variety of techniques of seeking to shape behaviour: legislation, licenses, circulars, regulations, registrations, administrative guidelines, codes of practice, industry standards, government incentives and private rights. And the choice of regulatory approaches and tools are related to factors such as the relevant context, stakeholders, and also the political and economic climate.[114]

[109] Department for Science, Innovation and Technology, "A Pro-Innovation Approach to AI Regulation: Government Response", updated 6 February 2024, https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response.
[110] Gikay, Asress Adimi, "Risks, Innovation, and Adaptability in the UK's Incrementalism Versus the European Union's Comprehensive Artificial Intelligence Regulation", *International Journal of Law and Information Technology* 32, no. 1 (2024), https://doi.org/10.1093/ijlit/eaae013; See also other contributions in this report.
[111] Black, Julia, "Critical Reflections on Regulation", *Australian Journal of Legal Philosophy* 27, (1) (2002): 1-27, 20.
[112] O'Mahony, Seamus, *Can Medicine be Cured; The Corruption of a Profession* (Apollo, 2019), 42.
[113] Brownsword, Roger, *Rights, Regulation and the Technological Revolution* (Oxford University Press, 2008), 285.
[114] Healy, Judith, *Improving Health Care Safety and Quality: Reluctant Regulators* (Ashgate, 2011).

In the UK, the four most relevant regulators are the Medicines and Healthcare products Regulatory Agency (MHRA),[115] who regulate medicines and medical devices, including software as a medical device, the National Institute for Health and Care Excellence (NICE),[116] who evaluate the clinical and cost effectiveness of health technologies and produce evidence-based guidance and advice for health, public health and social care, the Health Research Authority (HRA),[117] who protect the rights of patients, including by regulating the use of data collected within health and social care for research and product development and the Care Quality Commission (CQC),[118] who monitor, inspect and regulate services to make sure they meet fundamental standards of quality and safety and publish findings. The use of AI in healthcare will also raise issues for professional regulators such as the General Medical Council (GMC)[119] in discharging its functions around education, training and discipline, and relevant to the wider range organisations in the patient safety regulatory landscape in the UK National Health Service.[120]

The origins of the UK approach to regulating AI can be traced to two key documents: a House of Lords Select Committee on AI[121] and the Government National Strategy for AI.[122] Both reject a robust regulatory approach in favour of a pragmatic "watch and wait" approach. Further policy detail has emerged in the publication of the White Paper March 2023,[123] and a Government response to its public consultation in February 2024.[124] The tension between enabling innovation and responsibly regulating is readily apparent in the White Paper which:

> "details how we intend to support innovation while providing a framework to ensure risks are identified and addressed. However, a heavy-handed and rigid approach can stifle innovation and slow AI adoption. That is why we set out a proportionate and pro-innovation regulatory framework. Rather than target specific technologies, it focuses on the context in which AI is deployed. This enables us to take a balanced approach to weighing up the benefits versus the potential risks."[125]

---

[115] Medicines & Healthcare products Regulatory Agency, Accessed 25 March 2025, https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency
[116] National Institute for Health and Care Excellence, Accessed 25 March 2025, https://www.nice.org.uk/
[117] Health Research Authority, NHS, Accessed 25 March 2025, https://www.hra.nhs.uk/
[118] Care Quality Committee, Accessed 25 March 2025, https://www.cqc.org.uk/
[119] General Medical Council, Accessed 25 March 2025, https://www.gmc-uk.org/
[120] Oikonomou, Eirini, Carthey, Jane, Macrae, Carl, et al., "Patient Safety Regulation in the NHS: Mapping the Regulatory Landscape of Healthcare" *British Medical Journal Open* 9, no. 7 (2019): e028663. https://doi.org/10.1136/bmjopen-2018-028663
[121] UK Parliament, "UK Can Lead the Way on Ethical AI, Says Lords Committee", 2018, https://committees.parliament.uk/committee/376/artificial-intelligence-committee/news/94648/uk-can-lead-the-way-on-ethical-ai-says-lords-committee/
[122] GOV.UK, "National AI Strategy", Accessed 25 March 2025, https://www.gov.uk/government/publications/national-ai-strategy
[123] Office for Artificial Intelligence of the Department of Science, Innovation and Technology, "A Pro-Innovation Approach to AI Regulation", updated 3 August 2023, https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper
[124] Department for Science, Innovation and Technology, "A Pro-Innovation Approach to AI Regulation: Government Response", updated 6 February 2024, https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response
[125] Office for Artificial Intelligence of the Department of Science, Innovation and Technology, "A Pro-Innovation Approach to AI Regulation", updated 3 August 2023 https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper, p.2.

In short, the balance is skewed in favour of prioritising innovation at the expense of preventing risks to safety. There are no recommendations for an AI-specific regulator or new laws. Regrettably there is no detailed mapping of the currently regulatory regime, to understand what gaps there may be, leaving us unable to evaluate whether the status quo is fit for purpose.[126] In the absence of new laws or a specific regulator, the UK approach relies on soft compliance with the following cross sectoral principles:[127]

- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

These are less exacting than the OECD principles,[128] from where they emanate, for example, in omitting to refer to privacy, human rights, and societal wellbeing. Regulators *may* be under a statutory duty to have regard to these principles, though this represents a rather weak requirement even if implemented.

How can we summarise the strengths and weaknesses of this UK approach? The strengths are that it encourages innovation, allows a flexible and adaptive system of regulation, creates sector-specific guidelines, and enables regulatory clarity and simplicity. Nevertheless, the obvious disadvantages are the absence of enforceable statutory powers, imprecise principles, a disjointed regulatory landscape, and insufficient engagement with safety risks and ensuring effective redress mechanisms for AI related healthcare harm. In particular, safety risks appear to have been understated – for example, the White Paper claims that AI *"could"* pose a risk to safety, when the reality is that it *will* pose a risk.[129] There is no established definition of AI Safety[130] but a wealth of evidence that improving safety is challenging and requires carefully designed and sustained efforts.[131] The Government publications are also remarkably thin in terms of contestability and redress, and a striking contrast to the robust risk based system developed at EU level. The current UK approach of "keep calm and carry on" is

[126] See Charlesworth, Andrew, Fotheringham, Kit, Gavaghan, Colin, Sanchez-Graells, Albert and Torrible, Clare "Response to the UK's March 2023 White Paper 'A Pro-Innovation Approach to AI Regulation'", SSRN (2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4477368.
[127] Office for Artificial Intelligence of the Department of Science, Innovation and Technology, "A Pro-innovation Approach to AI Regulation", updated 3 August 2023, https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper at para 48.
[128] Organisation for Economic Co-operation and Development (OECD), "AI Principles", https://www.oecd.org/en/topics/sub-issues/ai-principles.html.
[129] Office for Artificial Intelligence of the Department of Science, Innovation and Technology, "A Pro-Innovation Approach to AI Regulation", updated 3 August 2023, https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper, p.27.
[130] Davies, Matt and Birtwistle, Michael "Regulating AI in the UK: Strengthening the UK's Proposals for the Benefit of People and Society", Ada Lovelace Institute (2023), https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/, p.21.
[131] Quick, Oliver, *Regulating Patient Safety: The End of Professional Dominance* (Cambridge University Press, 2017); Dixon-Woods, Mary "Why is Patient Safety so Hard? A Selective Review of Ethnographic Studies." *Journal of Health Services Research & Policy* 15, Suppl 1 (2010): 11-16, https://doi.org/10.1258/jhsrp.2009.009041.

complacent in terms of risks to safety and retains an uncertain legal framework which appears unable to provide adequate remedies for those harmed.[132]

To conclude, it remains unclear *how* this principles-based approach to regulation will be delivered. Crucially, how will these be developed into enforceable requirements? There is no coherent, deliverable regulatory model, and no independent oversight with the Government – through the Department for Science, Innovation and Technology[133] responsible for delivering central functions (co-ordination, monitoring, adapting). In addition, the "innovation first" requirement risks compromising regulatory independence.[134] For example, the MHRA sets out to "put patients first" yet this might be at risk with a Government mandate to put innovation first.

The lack of statutory approach means that the principles will be harder to deliver. There is a real risk that existing regulators are not empowered to act, and could remain open to public law review mechanisms if they act beyond their statutory powers.[135] The Government publications provide no steer on how to resolve conflicts when these principles inevitably conflict. The current approach is correctly described as one of deregulation[136] or unregulation.[137] Crucially, current mechanisms appear insufficient to ensure sufficient safety and adequate redress for AI related healthcare harm. Arguably, it is not possible to put innovation *and* safety first – there is always a trade off, and in deciding against designing a robust system of regulation to minimise the risk of harm, the current UK approach appears to put safety second.

[132] Smith, Helen and Fotheringham, Kit "Exploring Remedies for Defective Artificial Intelligence Aids in Clinical Decision-making in Post-Brexit England and Wales." *Medical Law International* 22, no. 1 (2022): 33-51, https://doi.org/10.1177/09685332221076124.
[133] Department for Science, Innovation and Technology, Accessed 25 March 2025, https://www.gov.uk/government/organisations/department-for-science-innovation-and-technology.
[134] Charlesworth, Andrew, Fotheringham, Kit, Gavaghan, Colin Sanchez-Graells, Albert and Torrible, Clare "Response to the UK's March 2023 White Paper 'A Pro-Innovation Approach to AI Regulation'", SSRN (2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4477368, p.5.
[135] Charlesworth, Andrew, Fotheringham, Kit, Gavaghan, Colin, Sanchez-Graells, Albert and Torrible, Clare "Response to the UK's March 2023 White Paper 'A Pro-Innovation Approach to AI Regulation'", SSRN (2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4477368, p.5.
[136] Roberts, Huw et al., "Artificial Intelligence Regulation in the United Kingdom: A Path to Good Governance and Global Leadership?" *Internet Policy Review* 12, no. 2 (2023), https://doi.org/10.14763/2023.2.1709.
[137] See Marsden, Christopher T, "Generative AI Regulation in UK", Chapter 44 in *Handbook on the Foundations and Regulation of Generative AI*, edited by Philipp Hacker, Sarah Hammer, Andreas Engel, Brent Mittelstadt (Oxford University Press, 2024).

Part 5: US – 1

Transcript of Presentation titled "Regulating AI/ML in U.S. Healthcare: Challenges, Opportunities, and FDA's Evolving Framework"

Presentation given by:
Prof Sara Gerke
Associate Professor of Law and Richard W. & Marie L. Corman Scholar,
Associate Professor, European Union Center,
University of Illinois Urbana-Champaign, United States

On 15 November 2024 at "Governing the Real World Application of Medical AI" Conference held by the Centre for Medical Ethics and Law (CMEL) of the University of Hong Kong

Transcript prepared by:
Ms Cordelia Chan

The following is a transcript of Prof Sara Gerke's presentation titled "Regulating AI/ML in US Healthcare: Challenges, Opportunities, and FDA's Evolving Framework" given on 15 November 2024 at the "Governing the Real World Application of Medical AI" Conference at the University of Hong Kong. This transcript has been lightly edited for clarity and up-to-date information.

Sara Gerke:

My presentation approaches health Artificial Intelligence (AI)/Machine Learning (ML)-based products from a "life cycle" perspective. This life cycle can be divided into three parts: the premarket, market, and postmarket stages.[138]

As the first step in the premarket stage, when developing a health AI/ML-based product, it is important to consider whether the product is a "medical device". If it is not a "medical device", the United States (US) Food and Drug Administration (FDA) may not have the authority to regulate it. The definition of "medical device" is set out in the US Federal Food, Drug, and Cosmetic Act (FDCA) section 201(h)(1), which says as follows:

> The term "device" (…) means an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is—
>
> (A) recognized in the official National Formulary, or the United States Pharmacopeia, or any supplement to them,
> (B) **intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man** or other animals, or
> (C) intended to affect the structure or any function of the body of man or other animals, **and which does not achieve its primary intended purposes through chemical action within or on the body of man** or other animals **and which is not dependent upon being metabolized for the achievement of its primary intended purposes**. The term "device" **does not include software functions excluded pursuant to section 520(o).**[139] (emphasis added)

[138] See Gerke, Sara, "Health AI for Good Rather Than Evil? The Need for a New Regulatory Framework for AI-Based Medical Devices", *Yale Journal of Health Policy, Law, and Ethics* 20, no. 2 (2021): 433-513.
[139] United States Federal Food, Drug, and Cosmetic Act, Section 201(h)(1); 21 U.S.C. § 321(h)(1).

This means that, in general, the term "medical device" includes software functions. Software can be divided into three types: (1) Software as a Medical Device ("SaMD"), (2) Software in a Medical Device ("SiMD"), and (3) Software utilized in the medical device's maintenance or manufacture.[140] I will focus on (1) SaMD and (2) SiMD. SaMD has been defined by the International Medical Device Regulators Forum as "software intended to be used for one or more medical purposes that perform these purposes **without being part of a hardware medical device**" (emphasis added).[141] SaMD is therefore a standalone software: the software itself is considered the "medical device".[142] SaMD can be operated with non-medical devices, such as laptops, smartwatches, smartphones, etc.[143] For example, Apple's electrocardiogram app is a SaMD.[144] In contrast, SiMD is software embedded in a medical device; there is both a software and a hardware component.[145] The software, in this case, should facilitate the medical device's function in some way.[146] A good example is software that aids in running an insulin pump.[147] The majority of AI tools on the market and being developed in the US are SaMD.[148]

It is important to realise that some software functions are not considered "medical devices" at all in the US. Under section 520(o) of the FDCA, there are five categories of non-device software functions: (1) Administrative support of health care facilities;[149] (2) Maintenance or encouragement of healthy lifestyles;[150] (3) Serving as electronic patient records;[151] (4) Transferring, storing, converting formats, or displaying data and results;[152] and (5) Clinical decision support software.[153] In particular, (5) Clinical decision support software is very relevant for health AI/ML-based products: it includes software functions that are intended for the purpose of "supporting or providing recommendations to a health care professional" and "enabling such healthcare professional to independently review the basis for such recommendations that such software presents".[154] The phrase "independently review the basis for such recommendations" is a vague term, which opens up many possibilities for interpretation. Some AI manufacturers may try to make sure their product does not fall under the definition of "medical device", so that it falls outside the FDA's purview.

As the second step in the premarket stage, if a health AI/ML-based product falls within the definition of a "medical device," the manufacturer should then consider the regulatory pathways it needs to go through, as well as the applicable requirements.[155] In the US, medical devices are categorised into

---

[140] See Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79-163, at 98. https://ssrn.com/abstract=4404252.
[141] International Medical Device Regulators Forum, "Software as a Medical Device (SaMD): Key Definition", 2013, https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf.
[142] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79-163, at 98-99. https://ssrn.com/abstract=4404252.
[143] *Id.* at 99.
[144] *Id.*
[145] *Id.*
[146] *Id.*
[147] *Id.*
[148] *Id.*
[149] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(A); 21 U.S.C. Section 360j(o)(1)(A).
[150] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(B); 21 U.S.C. Section 360j(o)(1)(B).
[151] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(C); 21 U.S.C. Section 360j(o)(1)(C).
[152] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(D); 21 U.S.C. Section 360j(o)(1)(D).
[153] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(E); 21 U.S.C. Section 360j(o)(1)(E).
[154] United States Federal Food, Drug, and Cosmetic Act, Section 520(o)(1)(E); 21 U.S.C. Section 360j(o)(1)(E).
[155] See Gerke, Sara, "Health AI for Good Rather Than Evil? The Need for a New Regulatory Framework for AI-Based Medical Devices", *Yale Journal of Health Policy, Law, and Ethics* 20, no. 2 (2021): 433; Gerke, Sara, Carmel Shachar,

three classes based on risk. Class I poses the lowest risk; Class II poses moderate risk; and Class III poses the highest risk.

There are four relevant types of premarket submissions. The first one is the so-called 510(k) (Premarket Notification). This pathway is applicable to Class I and Class II devices (subject to exemptions). The second type is the De Novo Classification Request. This pathway is for novel low-to-moderate risk devices. The third pathway is premarket approval ("PMA"), which is for Class III devices. The fourth pathway is the Humanitarian Device Exemption (HDE) for high-risk devices for rare conditions and diseases.[156]

The majority of AI/ML-based medical devices whose marketing has been permitted by the FDA went through the 510(k) pathway.[157] In the 510(k), sponsors only need to demonstrate that their medical devices are "substantially equivalent" to a legally marketed device. In other words, the 510(k) pathway does not usually require any clinical evidence.[158] This can be particularly worrisome for deep-learning AI. Deep learning uses a complex algorithm reasoning that makes it difficult, if not impossible, for humans to understand how the AI arrived at its output, making it non-interpretable.[159] The FDA has so far typically permitted the marketing of AI/ML-based medical devices with "locked algorithms".[160] The term is defined as "an algorithm that provides the same result each time the same input is applied to it and does not change with use".[161]

Of course, most AI algorithms are actually adaptive. This is perhaps their major strength, as they are able to continuously learn from new data.[162] In 2019, the FDA suggested in a discussion paper to implement a "Total Product Lifecycle (TPLC) Approach", so that the AI/ML-based medical devices are allowed to continuously improve.[163] Congress has recently provided the FDA with the authority needed to turn the plan into reality.[164] What AI/ML manufacturers can do now is to voluntarily submit a so-called "predetermined change control plan" with their premarket submission. If the FDA authorises such a plan, then they can make changes to the AI/ML-based medical device according to the plan without needing to undergo an otherwise required FDA review.[165]

Peter R. Chai, and I. Glenn Cohen, "Regulatory, Safety, and Privacy Concerns of Home Monitoring Technologies During COVID-19", *Nature Medicine* 26, no. 8 (2020): 1176-1182.

[156] See Gerke, Sara, Carmel Shachar, Peter R. Chai, and I. Glenn Cohen, "Regulatory, Safety, and Privacy Concerns of Home Monitoring Technologies During COVID-19", *Nature Medicine* 26, no. 8 (2020): 1176-1182.

[157] See United States Food and Drug Administration (FDA), Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices, https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.

[158] Gerke, Sara, "Health AI for Good Rather Than Evil? The Need for a New Regulatory Framework for AI-Based Medical Devices", *Yale Journal of Health Policy, Law, and Ethics* 20, no. 2 (2021): 433-513, at 474.

[159] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79, at 90.

[160] *Id.* at 91.

[161] United States Food and Drug Administration (FDA), "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", at 3, https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf.

[162] *Id.* Gerke, Sara, "Health AI for Good Rather Than Evil? The Need for a New Regulatory Framework for AI-Based Medical Devices", *Yale Journal of Health Policy, Law, and Ethics* 20, no. 2 (2021): 433-513, at 442.

[163] United States Food and Drug Administration (FDA), "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", https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf.

[164] United States Federal Food, Drug, and Cosmetic Act, Section 515C, 21 U.S.C. Section 316e–4.

[165] Gerke, Sara, "A Comprehensive Labeling Framework for Artificial Intelligence (AI)/Machine Learning (ML)-Based Medical Devices: From AI Facts Labels to a Front-Of-Package AI Labeling System — Lessons Learned from Food Labeling", *Emory Law Journal* (Vol. 74, 2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5113487. See United States Food and Drug Administration, "Marketing Submission Recommendations for a Predetermined Change

Another issue is the labelling of AI/ML-based devices. Currently, there are no labelling standards that are tailored to AI/ML-based medical devices.[166] There are only labelling requirements for medical devices generally, which are listed in Title 21 of the Code of Federal Regulations. These general labelling requirements do not work well for AI/ML-based medical devices. For example, according to one study, only 7 out of 161 US-marketed AI/ML-based medical devices provided publicly accessible race and ethnicity information; only 13 out of 161 disclosed gender information.[167] The majority of manufacturers also did not provide geographic breakdowns or detailed information on the validation data.[168] As this study confirms, the current labelling requirements in place for medical devices are insufficient for AI/ML-based medical devices.[169]

Without the creation of labelling standards that are tailored to AI/ML-based medical devices, many users will not receive sufficient information for their safe use, which could result in patients being harmed through biased care or pointless treatment.[170] My argument is that new labelling standards for AI/ML-based devices should include the following eleven key types of information:

(1) Model identifiers;
(2) Model type;
(3) Model characteristics;
(4) Indications for use;
(5) Validation and model performance;
(6) Details on the data sets;
(7) Preparation before use and application;
(8) Model limitations, warnings, and precautions;
(9) Alternative choices;
(10) Privacy and security; and
(11) Additional information.[171]

For example, for (5), the label should provide users with information on the validation and performance results. Information such as the model's cross-site performances is important so that users can assess the device's reliability. It is common for deep-learning models to be evaluated only at one clinical site. However, research has shown that these models can have weaker and worse performance across clinical sites.[172] As another example, for (6), details about the datasets, such as the general ethnicity breakdown, are very much necessary.[173]

Control Plan for Artificial Intelligence-Enabled Device Software Functions", 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence.

[166] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79, at 121-148.

[167] See Ross, Casey, "Explore STAT's Database of FDA-Cleared AI Tools", *STAT*, 3 February 2021. https://www.statnews.com/2021/02/03/fda-artificial-intelligence-clearance-products.

[168] *Id.*

[169] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79, at 143.

[170] *Id.*

[171] *Id.* at 148-49.

[172] *Id.* Wu, Eric, Kevin Wu, Roxana Denshjou, David Ouyang, Daniel E. Ho and James Zou, "How Medical AI Devices Are Evaluated: Limitations and Recommendations from an Analysis of FDA Approvals", *Nature Medicine* 27 (2021): 582-584, at 583.

[173] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79, at 154.

The list is non-exhaustive, but it should help regulators such as the FDA to begin the overdue creation of labelling standards that are tailored to AI/ML-based medical devices.[174] During that process, the FDA should include all stakeholders, such as patient and consumer representatives.[175] Besides general labelling standards for AI/ML-based medical devices, the FDA and other regulators also need to consider additional labelling standards for specific users and/or types of AI.[176] Regulators need to do a better job of educating users about the risks of AI/ML-based medical devices. With the right design, labels could help achieve that goal. For example, an "eye-popping" design, which is similar to a nutrition facts label, could be helpful, as the design is both familiar to users and enables them to get a quick summary of the key information.[177]

At the market stage, once a premarket submission has been successfully completed, the FDA may permit marketing of the AI/ML-based medical device. So far, the FDA has authorized over 1000 AI/ML-based medical devices.[178] Most of these devices are intended for use in radiology, followed by cardiology and neurology.[179] There are also marketed AI/ML-based medical devices in other specialties, such as ophthalmology and gastroenterology.[180] Out of all marketed AI/ML-based medical devices, the majority (980) received 510(k) clearance,[181] which, as mentioned, only requires that the device is shown to be "substantially equivalent" to a legally marketed device, mostly with no clinical evidence needed. 32 of the AI/ML-based medical devices went through the De Novo process, whereas only 4 went through the premarket approval (PMA) pathway, which is the strictest regulatory pathway.[182]

An example of a marketed AI/ML-based medical device is LumineticsCore™ (formerly IDx-DR). It is a medical device which uses AI to detect diabetic retinopathy in patients. This is the first device to be considered autonomous, as its decision is not reviewed by a human (here, the primary care physician).[183] It has been marketed since April 2018 in several sites across the US.[184]

At the postmarket stage, once an AI/ML-based medical device is launched onto the market, continuous risk monitoring becomes crucial. In particular, it will be a necessity for adaptive algorithms. Regulators such as the FDA need to shift away from a "product view" towards a "system view" perspective: they "need to widen their scope from evaluating medical AI/ML-based products to assessing systems".[185] This view is crucial "to maximizing the safety and efficacy of AI/ML in health care".[186] However, it also poses a challenge as regulators such as the FDA are accustomed to regulating products and not systems.[187] Nevertheless, this shift is crucial as AI/ML-based medical

---

[174] *Id.* at 149.
[175] *Id.*
[176] *Id.*
[177] *Id.* at 158-59, 163.
[178] United States Food and Drug Administration (FDA), "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices", https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.
[179] Gerke, Sara, "A Comprehensive Labeling Framework for Artificial Intelligence (AI)/Machine Learning (ML)-Based Medical Devices: From AI Facts Labels to a Front-Of-Package AI Labeling System — Lessons Learned from Food Labeling", *Emory Law Journal* (Vol. 74, 2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5113487.
[180] *Id.*
[181] *Id.*
[182] *Id.*
[183] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79-163, at 89.
[184] Digital Diagnostics, "LumineticsCore™", https://www.digitaldiagnostics.com/products/eye-disease/lumineticscore/.
[185] Gerke, Sara, Boris Babic, Theodoros Evgeniou, and I. Glenn Cohen, "The Need for a System View to Regulate Artificial Intelligence/Machine Learning-based Software as Medical Device", *npj Digital Medicine* 3 (2020): 53. https://doi.org/10.1038/s41746-020-0262-2.
[186] *Id.*
[187] *Id.*

devices are usually one part of a larger system featuring multiple kinds of human involvement.[188] For example, healthcare teams may input the data, the physician may rely on the AI's recommendation, and insurers may decide if they want to reimburse the treatment; the entire system must be assessed.[189] Consequently, it might be helpful to reframe what hospitals are actually doing: they are not "buying" an AI/ML tool, they are "hiring" one.[190] In the same way that cognitive testing of physicians will not tell hospitals how they will do when put into a pre-existing team, the same is also true for AIs; they need to be continuously evaluated.[191] The "system view" is particularly important for generative AI tools such as Large Language Models. It will be even more imperative in the future to think about how we can effectively assess these systems.

[188] *Id.*
[189] *Id.*
[190] *Id.*
[191] *Id.*

Part 5: US – 2

Managing Failure: How Should We Govern Problems Arising from Medical AI/ML Devices?

The following is a summary of a presentation given by *Prof Boris Babic on 15 November 2024 at the "Governing the Real World Application of Medical AI" Conference at the University of Hong Kong. The title of the presentation was "Managing Failure: How Should We Govern Problems Arising from Medical AI/ML devices?". This summary was prepared by Ms Cordelia Chan.

*Prof Boris Babic
Associate Professor, HKU Musketeers Foundation Institute of Data Science and
Department of Philosophy
Associate Professor (by courtesy), Faculty of Law
The University of Hong Kong, Hong Kong SAR, China

Prof Boris Babic's joint project (the authors' project) seeks to understand what happens to medical devices from a regulatory perspective once they have been brought to the market in the United States (US). Using data from the US Food and Drug Administration (FDA) 510(k) database, merged with the Manufacturer and User Facility Device Experience (MAUDE) adverse event reports, it examines the limits of adverse event reporting for Artificial Intelligence (AI) /Machine Learning (ML) devices and makes suggestions for how to improve the substantive regulatory tracking system for AI/ML devices once they are on the market.[192]

Regulatory Background

The authors' project begins by describing the regulatory background in the United States. According to the authors, modern medical device regulation began with the 1976 Medical Device Amendments (MDA) to the 1938 Federal Food, Drug and Cosmetic Act (FDCA). The authors' project focuses in particular on post-market surveillance of legally marketed medical devices – i.e., recalls and adverse event reports. Currently, there are too few recalls of AI/ML devices to draw any meaningful conclusions ($n < 10$). Accordingly, the project is limited to adverse event reporting in particular.

The FDA's Medical Device Reporting (MDR) system is one of the central post-market surveillance tools for managing medical device related adverse events. The core reporting requirements are laid out in the Medical Device Reporting Regulation (21 CFR, Part 803), which was published on 11 December 1995 (60 FR 63578). It is authorised by Section 519 of the FDCA.

The authors also note that the MDR requires manufacturers, user facilities, and importers of legally marketed medical devices to submit reports of certain adverse events involving their medical devices (21 CFR 803.10(a)-(c)). The FDA refers to these as MDR reportable events. MDR reportable events for device manufacturers include reports of death, serious injury, or device malfunction. A malfunction is reportable only if it would be likely to cause death or serious injury if the malfunction were to recur (21 CFR 803.3 and 21 CFR 803.50).

[192] Babic, Boris, I. Glenn Cohen, Ariel Dora Stern, Yiwen Li, and Mellisa Ouellet, "A General Framework for Governing Marketed AI/ML Medical Devices", *npj Digital Medicine* 8 (2025): 328. https://doi.org/10.1038/s41746-025-01717-9.

Data

According to the authors, the FDA uses a tiered risk system (Class I-III) for approving medical devices. Most AI/ML devices have been cleared under Moderate Risk (Class II). Moderate risk (Class II) devices are typically regulated through a process called Premarket Notification or, more often, the 510(k) process. This process requires a device to demonstrate substantial equivalence with one or more already legally marketed devices. Class II devices that do not have a legally marketed predicate device can use a De Novo Classification request.

The FDA maintains a comprehensive 510(k) database. The authors' project starts with the FDA's downloadable 510(k) files, restricting to medical devices approved from 2010 through 2023. This includes both Class I and II Premarket Notification (510(k)) devices and De Novo classification requests for low to moderate risk medical devices. This dataset represents 98% of all FDA device market authorisations over the same period. The data is then filtered for AI/ML-based medical devices. In some cases, a comparison is drawn between AI/ML-based and non-AI/ML-based medical devices.

The authors' project calculates the number of adverse events within specific timeframes post-approval (3, 6, 9, 12, 24 months) using the FDA's Medical Device Reporting (MDR) system. The final dataset comprises 823 unique 510(k)-cleared AI/ML devices that could be linked to a total of 943 subsequent adverse events reported ("MDRs" for short). This dataset tracks 54 features related to the reported events and device manufacturers. For example, features include the type of event, the setting where it occurred, the manufacturer and product it was associated with, and so forth. Within the 943 linked MDRs, there are 20 unique product codes.

Analysis

Across the final dataset, three product codes account for an overwhelming number of adverse events. They are the Mass Spectrometry Microbial Identification System (PEX), Dario (NBW), and HeartFlow (PJA). Most MDRs associated with the Mass Spectrometry Microbial Identification System (PEX) are for misidentifications of microorganisms. Prof Babic focused on HeartFlow (PJA) and Dario (NBW) in his presentation.

Based on the MDRs, the main issue for HeartFlow (PJA) is false negatives. HeartFlow is a medical technology company that was established in 2009. HeartFlow Analysis is a coronary physiologic simulation software used in hospitals for the diagnosis and management of coronary artery disease. It uses CT scans to create a 3D model of the patient's arteries and calculates Fractional Flow Reserve derived from CT scans (FFR CT), a mathematically derived quantity that indicates the extent of coronary artery disease. All MDRs of PJA were reported as false-negative results associated with the HeartFlow Analysis – i.e. indications of sufficiently high blood flow when blockage should have been identified. The source of/reason for error is hard to identify. Some MDRs attribute the problem to analyst error.

The main issue for Dario (NBW) is false positives. Dario is a direct-to-consumer app-based blood glucose monitoring system. Most reports involve high blood glucose readings that are false positives. Often these lead patients to seek additional medical services. It is hard to know what happens after or to diagnose precisely why the problem occurred. For example, a customer was contacted on numerous occasions to obtain additional details regarding the emergency room visit and to request the meter back for investigation. The customer did not respond to follow up attempts. (Event date: 20 February 2018.)

Missing Data and Misclassification

One major problem is missing data. "Event Location" is entirely missing in all MDRs in the sample. 73% of the reports lack information about whether the reporter was a health professional. Event Date is missing in 32% of the reports. Reporter Occupation is absent in 30% of the cases.

Importantly, the extent of missing data was significantly higher in the AI/ML sample compared to traditional medical devices. For instance, information about whether the reporter was a health professional was missing 73% of the time in the AI/ML sample, but only 43% in the overall sample.

In addition, data is also often misclassified. Adverse event reports are generally classified into three categories: malfunction, injury, and death. There are very few deaths: only 2 reported for NBW; and the deaths do not seem associated with the device, according to the qualitative description. The vast majority of the adverse event reports of AI/ML medical devices within the dataset of the authors' project was classified as a "malfunction". However, when the data is read qualitatively, the "malfunctions" are quoted as "analyst errors". Regression analyses on how safety is affected by location and event type cannot be carried out as much of the data is either missing or misclassified.

How to Improve

As a result, the data provides limited insight into how much risk AI/ML devices present once they are brought to the market. The FDA's tracking of adverse event reports is of limited usefulness, since the data is static and incomplete; and the people who report adverse events lack incentive to provide an accurate report. For example, a company would prefer to categorise a malfunction as an "analyst error", rather than a problem with its device. It is therefore difficult to know where the mistakes are coming from. In addition, there could potentially be a problem of over-reporting. The companies reporting a higher number of adverse events could just be more diligent with reporting than others.

There needs to be improvement in reporting characteristics. As a starting point, it is helpful to consider three issues previously identified by Boris Babic *et al*[193] and Sara Gerke *et al*:[194] concept drift, covariate shift, and algorithmic stability. These are important safety dimensions of AI/ML products that are not accounted for in the FDA's post-market regulatory system.

Concept drift concerns true joint feature/label distribution. The distribution that the algorithm is trying to approximate changes slightly over time. The relationship between inputs and outputs changes slightly as the AI/ML system is deployed. For example, an AI/ML-based admission triage tool which is tested on a sample corresponding to a general healthcare environment may then be deployed in managing intensive care unit (ICU) queues. The initially high specificity/sensitivity can deteriorate quickly after concept drift.

Covariate shift refers to a situation in which the feature distribution alone changes. This can occur if the training or early use data is not representative, but it can happen for other reasons as well. As an extreme case, consider a situation where an AI/ML medical device is trained on a sample of (high-risk for diabetes) obese young men and then applied to a sample of (high-risk for diabetes) non-obese older men. The distribution of diabetes among both groups could well be the same (i.e. the label distribution does not change), but the feature distribution is very different (obesity vs. age). This is a

[193] Babic, Boris, Sara Gerke, Theodoros Evgeniou, and I. Glenn Cohen, "Algorithms on Regulatory Lockdown in Medicine: Prioritize Risk Monitoring to Address the 'Update Problem'", *Science* 366, no. 6470 (2019): 1202-1204. https://doi.org/10.1126/science.aay9547.

[194] Gerke, Sara, Boris Babic, Theodoros Evgeniou, and I. Glenn Cohen, "The Need for A System View to Regulate Artificial Intelligence/Machine Learning-Based Software as Medical Device", *npj Digital Medicine* 3 (2020): 53. https://doi.org/10.1038/s41746-020-0262-2.

case of covariate shift without concept drift, and indeed without lack of representativeness in the data. Only the feature distribution itself is changing. Covariate shift can deteriorate a predictive diagnostic system very quickly.

For algorithmic stability, the Lipschitz Property requires that for a given distance between two observations in the feature space, their distance should be similarly bounded in the output space. Distance needs to be formally represented in a suitable measure space, keeping in mind that the dimensionality of the input space is different from the dimensionality of the label space. This is similar to the "treat like cases alike" desideratum in law; it is particularly important in diagnostics and resource allocation.

To address the problem, reporting requirements could be modified to require, for instance, that for every AI/ML medical device, manufacturers must flag both when training data is significantly updated and when deployment conditions are substantially amended. Manufacturers can also be required to, for example, produce a quarterly update reflecting changes in feature distribution, label distribution, and stability.

A more radical change would be to move beyond database reporting. Regulators may want to consider a different public health governance regime altogether. Sara Gerke argues in favour of nutrition label style reporting for AI/ML medical devices.[195] Such labels could reflect issues arising in feature and label distribution changes in the product's development. This could be combined with human factors testing and a system of continuous oversight.

[195] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, no. 1 (2023): 79-163. https://ssrn.com/abstract=4404252.

Part 6: Middle East

Regulating the "True" Lifecycle of Healthcare AI: Considerations from Research to Deployment

Dr Barry Solaiman
Associate Dean for Academic Affairs & Assistant Professor of Law, College of Law, Hamad Bin Khalifa University,
Qatar, Doha

The proliferation of discussions surrounding the governance of Artificial Intelligence (AI) compels us to critically examine the most appropriate means of regulation and the extent to which governance should intervene. The intersection of "Health, AI, and Law" has developed into a distinct field, with potential regulatory approaches spanning a broad spectrum. [196] Lawyers and policymakers have been navigating the path from "AI to Law," identifying novel challenges posed by AI in healthcare and assessing the limitations of existing laws to address these challenges. [197] In the interim, soft laws—non-binding guidelines or frameworks—have sought to fill a legal void while the development of enforceable, hard laws lags behind. Scholars are attempting to address matters surrounding the explainability of AI and informed consent, the impact of AI on the doctor-patient relationship, and the effect of AI in bespoke areas of healthcare such as mental health or long-term care, amongst others.[198] For the foreseeable future, it is likely that a hybrid model of hard and soft law will remain the dominant paradigm. Against this backdrop, the question arises: how should governance in this space evolve?

The current approach towards governance by the European Union (EU) under the Artificial Intelligence Act (EU AI Act) arguably fails to build trust in healthcare.[199] Instead, any regulatory system should contend with a "True Lifecycle Approach" (TLA) towards governing AI, which can be conceptualised in three core phases. [200] The first is the research and development (R&D) phase, encompassing the initial conception of an AI system. This includes coding the algorithm, training and validating the model, and deploying the system. Globally, there remains a significant gap in specific regulatory approaches to guide this phase in healthcare research. The second phase involves embedding AI into a medical device, where regulatory approvals come into play. Whilst a few countries have made strides in introducing AI-specific medical device guidelines, there is no unified or consistent approach across jurisdictions. The third phase pertains to the use of AI in clinical practice, whether for administrative or clinical purposes. Here, a pronounced regulatory chasm exists globally.

---

[196] Solaiman, Barry and I. Glenn Cohen, "A Framework for Health, AI and the Law," in *Research Handbook on Health, AI and the Law*, ed. Barry Solaiman and I. Glenn Cohen, p. 1-19 (Edward Elgar, 2024), https://doi.org/10.4337/9781802205657.ch01.

[197] Solaiman, Barry, "From 'AI to Law' in Healthcare: The Proliferation of Global Guidelines in a Void of Legal Uncertainty," *Medicine and Law* 42, no. 2 (2023): 391-406, https://ssrn.com/abstract=4519627.

[198] Solaiman, Barry, and Mark G. Bloom. "AI, Explainability, and Safeguarding Patient Safety in Europe." in *The Future of Medical Device Regulation: Innovation and Protection* pp. 91-102, https://doi.org/10.1017/9781108975452.008; Solaiman, Barry, and Abeer Malik. "Regulating Algorithmic Care in the European Union: Evolving Doctor–Patient Models Through the Artificial Intelligence Act (AI-Act) and the Liability Directives." *Medical Law Review* 33.1 (2025): fwae033, https://doi.org/10.1093/medlaw/fwae033; Solaiman, Barry. "Generative Artificial Intelligence (GenAI) and Decision-making: Legal & Ethical Hurdles for Implementation in Mental Health." *International Journal of Law and Psychiatry* 97 (2024), https://doi.org/10.1016/j.ijlp.2024.102028; Solaiman, Barry. "Legal and Ethical Considerations of Artificial Intelligence for Residents in Post-Acute and Long-Term Care." *Journal of the American Medical Directors Association* 25.9 (2024), https://doi.org/10.1016/j.jamda.2024.105105.

[199] Solaiman, Barry. "The European Union's Artificial Intelligence Act and Trust: Towards an AI Bill of Rights in Healthcare?." *Law, Innovation and Technology* 17.1 (2025): 318-334, https://doi.org/10.1080/17579961.2025.2469986.

[200] The author first proposed the TLA in Solaiman, Barry and others, "A 'True Lifecycle Approach' Towards Governing Healthcare-AI with the GCC as a Global Governance Model" in *npj Digital Medicine* (forthcoming 2025).

While these phases form the backbone of AI's journey, they do not encompass every aspect. Nevertheless, they provide a useful framework for regulatory consideration. Despite the lack of a unified global governance approach, developments in three neighbouring Gulf Cooperation Council (GCC) countries—Qatar, Saudi Arabia, and the United Arab Emirates (UAE)—may offer valuable insights.[201] Qatar focuses on the R&D phase, Saudi Arabia is a leader in the regulation of AI-based medical devices, and the UAE provides governance structures for clinical implementation. This approach presents a "True Lifecycle Approach" towards the governance of AI in healthcare for consideration by regulators globally. Indeed, a combined analysis of these efforts reveals the potential to address the lifecycle regulatory conundrum for AI in healthcare.

Phase 1: Qatar Pioneering AI Guidelines for Healthcare Research

In Qatar, a comprehensive research grant, completed in 2024 through Hamad Bin Khalifa University (HBKU), resulted in the creation of detailed guidelines and a draft certification process for AI in healthcare research. [202] These guidelines are notable for their technical precision, requiring information on the development, external validation, and deployment plans of AI models. Uniquely, they also integrate local demographic considerations, addressing nationality, race, religion, and other contextual factors to ensure that AI systems are tailored to Qatar and the broader Middle East region.

The guidelines were developed by a multidisciplinary team of experts from law, Islamic bioethics, science, biomedicine, and healthcare, many of whom were actively involved in AI research. Importantly, the Ministry of Public Health (MOPH) served as official advisers throughout the project, ensuring alignment with national policy priorities. Feedback from local expert roundtables refined the guidelines, which were finalised alongside a mock website to facilitate future implementation and testing.

As Qatar prepares for broader implementation, critical governance questions remain. Should these guidelines evolve into enforceable hard law, or remain as flexible soft law? How should they integrate into institutional review board (IRB) processes, or should they stand as an independent framework? Should a dedicated standing committee oversee compliance, and if so, what expertise should it encompass? Addressing these questions will shape Qatar's contribution to regulating AI in healthcare research, potentially offering a replicable model for other jurisdictions.

Phase 2: Saudi Arabia Setting Standards for AI-Based Medical Devices

Saudi Arabia's *MDS-G010: Guidance on Artificial Intelligence and Machine Learning Technologies Based Medical Devices*[203] demonstrates a distinctive approach to the regulation of AI-based medical devices. Unlike jurisdictions that adapt general medical device standards to AI, the Saudi Food and Drug Authority (SFDA) explicitly addresses the unique lifecycle of AI/ Machine Learning (ML) technologies. This includes requirements for transparency, data governance, and algorithmic accountability, as well as continuous monitoring and post-market surveillance to ensure ongoing safety and effectiveness.

---

[201] Solaiman, Barry, Ayesha Bashir, and Fama Dieng, "Regulating AI in Health in the Middle East: Case Studies from Qatar, Saudi Arabia, and the United Arab Emirates," in *Research Handbook on Health, AI and the Law*, ed. Barry Solaiman and I. Glenn Cohen, p. 332-354, (Edward Elgar, 2024), https://doi.org/10.4337/9781802205657.ch19.

[202] Solaiman, Barry, and others, "Research Guidelines for Healthcare AI Development" Version 1.0. Hamad Bin Khalifa University; April 2025. Developed under the research grant "Artificial Intelligence for Precision Medicine & Health Technologies: Developing a Regulatory Framework for Qatar and the Middle East" (HBKU-SRO-TGA-VPR-TG01-001), http://dx.doi.org/10.13140/RG.2.2.10590.14402.

[203] Saudi Food and Drug Authority, "Guidance on Artificial Intelligence (AI) and Machine Learning (ML) Technologies Based Medical Devices (MDS-G010)", Version 1.0, November 29, 2022, https://www.sfda.gov.sa/sites/default/files/2023-01/MDS-G010ML.pdf.

What sets Saudi Arabia apart is its emphasis on rigorous validation of AI models in culturally and regionally specific healthcare contexts.[204] This approach ensures that algorithms are not only technically robust but also ethically aligned with local values. For example, the SFDA mandates that AI models account for diverse populations and regional healthcare needs, positioning Saudi Arabia as a leader in AI/ML device regulation. Its framework could serve as a template for nations seeking to establish AI-specific standards that balance innovation with ethical responsibility.

Phase 3: The UAE and Governance at the Clinical Implementation Stage

In the UAE, two policies stand out for addressing the governance of AI at the clinical implementation stage. Abu Dhabi's *Policy on Use of Artificial Intelligence in the Healthcare Sector* (2018),[205] introduced by the Department of Health, integrates AI into the healthcare system with a dual focus on operational enhancements and patient safety. This policy is special for mandating alignment with international standards while tailoring AI applications to local cultural and regulatory contexts. By prioritising compatibility with local norms, Abu Dhabi demonstrates a commitment to ensuring that AI technologies serve the specific needs of its population.

Dubai's *Artificial Intelligence in the Health Sector Policy* (2021),[206] issued by the Dubai Health Authority, takes a different but complementary approach. This policy embeds ethical considerations into the regulatory framework, emphasising transparency, accountability, and the delineation of roles for stakeholders. Few other global jurisdictions have made such explicit commitments to ethical governance. Together, Abu Dhabi and Dubai provide a forward-thinking blueprint for balancing innovation with robust oversight in healthcare.

A Combined Approach: Lessons from the GCC

The regulatory efforts in Qatar, Saudi Arabia, and the UAE underscore the importance of addressing AI governance across its lifecycle. Qatar's detailed R&D guidelines, Saudi Arabia's comprehensive framework for medical devices, and the UAE's policies for clinical implementation collectively highlight how regional contexts can shape innovative regulatory approaches. A combined approach that integrates these efforts could offer a holistic model for governing AI in healthcare, addressing gaps in each phase of the lifecycle. For example, Qatar's guidelines could inform the development of robust approval processes, while Saudi Arabia's emphasis on post-market surveillance could enhance the UAE's clinical implementation frameworks.

Conclusion

The governance of AI in healthcare is an evolving field, fraught with complexity and opportunity. The GCC's efforts highlight the potential for regionally tailored approaches to set global benchmarks. Qatar's pioneering work in R&D, Saudi Arabia's leadership in medical device regulation, and the UAE's focus on clinical implementation collectively offer a roadmap for addressing the AI lifecycle. By integrating these efforts into a cohesive framework, the GCC approach can offer a model for developing comprehensive and culturally sensitive governance structures. This combined approach

[204] For a detailed analysis of the MDS-G010 that offers a pinpoint analysis of its provisions, see Solaiman, Barry, "Regulating AI-Based Medical Devices in Saudi Arabia: New Legal Paradigms in an Evolving Global Legal Order." *Asian Bioethics Review* 16 (2024): 373-389, https://doi.org/10.1007/s41649-024-00285-6.
[205] Department of Health Abu Dhabi, "Policy on Use of Artificial Intelligence (AI) in the Healthcare Sector of the Emirate of Abu Dhabi", 2018, https://www.doh.gov.ae/-/media/E9C1470A575146B18015DEBE57E47F8D.ashx.
[206] Dubai Health Authority, "Artificial Intelligence in the Health Sector Policy", 2021, https://dha.gov.ae/uploads/082022/Artificial%20Intelligence%20Policy20228457.pdf.

not only addresses immediate regulatory needs but also spurs further debate and research into crafting the most effective pathways for AI governance in healthcare.

Part 7: Not Jurisdiction-specific – 1

The Case for Discretionary Principles in Determining the Usage of AI in Medical and Healthcare Settings

Prof Brian Wong
Assistant Professor, Department of Philosophy, The University of Hong Kong,
Hong Kong, China

Usage of *Artificial Intelligence* (AI) is increasingly ubiquitous in medical and healthcare settings. From diagnostic models to health insurance algorithms, from systems aimed at identifying likely treatments and offering prognostic predictions to those tasked with allocating scarce resources in the face of medical emergencies[207], it is apparent that AI is accruing substantial significance. Yet theorisation of the regulation and deployment of, and accountability for, AI usage remains heavily finite. Accounts in existing literature tend to adopt a top-down method in grappling with key issues and are either excessively deferential to foundational principles of medical ethics without considering challenges in their applicability, or do not pay sufficient attention to the *sui generis* nature of the doctor-patient relationship by presuming that the relevant discussions can be directly transplanted from broader discussions on AI management and accountability at large.[208]

What is needed, especially on the part of policymakers – given their standing as actors tasked with legislating and formalising regulatory principles and ordinances – is a framework that allows for *justifiable discretion.* Such a framework should allow for actors to be cognisant of circumstantial sensitivities whilst drawing upon a robust and comprehensive system of normative principles in their prescriptions.

In meeting this demand, this paper begins by considering the four most commonly cited *core principles* of medical ethics.[209] Beneficence is best defined as acting in the best interest of the individual. Non-maleficence is best construed as the avoidance and refraining from perpetration of harm to the individual. Autonomy requires that the individual be given the freedom to choose freely – with stipulations that consent must be well-informed and not extracted through coercion or deception. Justice, which is perhaps the most amorphous term to define clearly, can be broadly interpreted as the requirement that all individuals are treated equally and equitably. Whilst such principles are most commonly applied to the set of *patients* in relation to their care providers (e.g. doctors, nurses, and other medical personnel), it should be noted that such considerations can also be extended to other stakeholders within the medical setting. Doctors should not behave in a way that puts their colleagues at risk, on grounds of non-maleficence. Nurses should have their autonomy

[207] Alowais, Shuroug A, Sahar S Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I Alshaya, Sumaya N Almohareb, Atheer Aldairem, Mohammed Alrashed, Khalid Bin Saleh, Hisham A Badreldin, Majed S Al Yami, Shmeylan Al Harbi, and Abdulkareem M Albekairy, "Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice," *BMC Medical Education* 23, no.1 (2023): 689, ; Alam, Ashrafe and Victor R. Prybutok, "Use of Responsible Artificial Intelligence to Predict Health Insurance Claims in the USA Using Machine Learning Algorithms," *Exploration of Digital Health Technologies* 2, (2024): 30 – 45.
[208] Farhud, Dariush D. and Shaghayegh Zokaei, "Ethical Issues of Artificial Intelligence in Medicine and Healthcare," *Iranian Journal of Public Health* 50, no.11 (2021): 1 – 5; Morley, Jessica, Caio C.V. Machado, Christopher Burr, Josh Cowls, Indra Joshi, Mariarosaria Taddeo, and Luciano Floridi, "The Ethics of AI in Health Care: A Mapping Review," *Social Science & Medicine* 260 (2020): 113 – 172.
[209] Varkey, Basil, "Principles of Clinical Ethics and Their Application to Practice," *Medical Principles and Practice: International Journal of the Kuwait University, Health Science Centre* 30, no.1 (2021): 17 – 28; Gillon, Raanan, "Ethics Needs Principles—Four Can Encompass the Rest—and Respect for Autonomy Should be 'First Among Equals'," *Journal of Medical Ethics* 29, no.5 (2003): 307 – 312.

enshrined and protected by their workplace, so long as their autonomy does not entail their undermining the interests of other actors.

Yet the rise of AI – especially in its current forms – risks contravening these principles. An excess reliance upon AI models as the basis for diagnosis, prognosis, and prescription would run into the serious obstacle of algorithmic bias and injustice. If the training data for AI omits certain demographics, provides incomplete or skewed representation, or disproportionately represents any group or groups over other group(s) (e.g. cis-heterosexual men to the exclusion of individuals of other gender identities and sexual orientations), then the AI could well become biased as a result.[210] Studies have found that in the US, biased algorithms have required persons of colour to be significantly more ill – exhibiting far more severe symptoms – so as to receive the same diagnosis and treatment as their white counterparts.[211] A 2019 study concluded that a more AI-dominated bio-medical healthcare system may have "adverse impacts on individuals with complex needs", failing to address the intersectional and multi-faceted nature of medical needs by patients that cannot be reduced into simplistic archetypes. A mixture of incomplete or skewed data significantly constrains the *input quality* of AI model and yields undue violations of the *Beneficence* and *Non-maleficence* criteria.

Furthermore, an over-deference to assessments and recommendations by AI models by physicians and medical practitioners may also undermine another set of interests held specifically by patients and their significant others – that is, their *epistemic agency.* The veneer of objectivity and gravitas associated with machine learning-based models could thus lead to practitioners dismissing the testimonies[212] of patients and discarding their self-reported experiences, which may be harder to precisely quantify and thus appear lacking in credibility as compared with AI-yielded conclusions. Additionally, patients may wish to opt out of being treated with AI-based diagnoses, prognoses, and prescriptions. What the AI construes to be 'best' may not in fact be what the patient views as 'best', let alone what is ultimately 'best' – this discrepancy in turn gives rise to a broader concern of undue paternalism. The inability of medical systems to accommodate the wishes of patients and their associates, thus amounts to an instance of epistemic injustice, which undermines the *Autonomy* and *Justice* criteria.

Neither of these issues amounts to intractable problems when it comes to AI usage in medical contexts. Nor will these issues be completely erased in a counterfactual without AI usage. Yet we should and can aim to do better. Neither complete prohibition on nor unfettered deployment of AI can be a viable path forward.

This paper suggests that in generating regulations and protocol governing AI usage, as well as conceptualising broader policies on AI adoption in medical contexts, policymakers should devise a set of *discretionary principles* that strive to satisfy four key objectives:

1. Ensuring that infringements of the above four principles (Beneficence, Non-Maleficence, Autonomy, and Justice) are mitigated, and preventatively minimised;

[210] Krasniansky, Adriana, "Understanding Racial Bias in Medical AI Training Data," *The Blog of the Petrie-Flom Center at Harvard Law School*, October 29, 2019, https://petrieflom.law.harvard.edu/2019/10/29/understanding-racial-bias-in-medical-ai-training-data/; Stetler, Carrie, "AI Algorithms Used in Healthcare Can Perpetuate Bias," *Rutgers University - Newark Research & Innovation,* November 14, 2024, https://www.newark.rutgers.edu/news/ai-algorithms-used-healthcare-can-perpetuate-bias.

[211] Backman, Isabella, "Eliminating Racial Bias in Health Care AI: Expert Panel Offers Guidelines," *Yale School of Medicine*, December 21, 2023, https://medicine.yale.edu/news-article/eliminating-racial-bias-in-health-care-ai-expert-panel-offers-guidelines/.

[212] Fricker, Miranda, "Testimonial Injustice," in *Epistemic Injustice: Power and the Ethics of Knowing,* 9 – 29 (Online: Oxford Academic, 2007).

2. If the infringement of one or more of these four principles is inevitable, then there must be either a) sufficient compensation for such infringement that can be viably delivered to the affected party or b) reasonable justification for the tradeoff;
3. Ensuring that the positive interests stipulated by each of these principles are proactively maximised;
4. Providing for sufficient space for meaningful deliberation, quality debate, and viable disagreement between disparate stakeholders, to serve as check-and-balance against the formal powers of policymakers.

These principles are loosely arranged in a descending order of importance. The first and foremost consideration is ensuring that individual stakeholders are not deleteriously impacted by the incorporation of AI. Should infringements occur, they must be duly justifiable or compensated, such that they do not become violations.[213]

Somewhat lower down in the hierarchy of priorities come the advancement of *upsides*, as well as the maintenance of sufficient room for deliberation, debate, and critical scrutiny and reflexivity on the part of all involved actors. This is much needed to prevent epistemic violations and injustices within the doctor-patient, as well as more general, relationship within the medical context.

The discretionary principles legislated by lawmakers in accordance with the above, would in turn give rise to what I term *discretionary duties* – morally demanding prescriptions weighty in force yet flexible with regards to their exact contents. Such duties are correlated broadly with *pro tanto rights* held by individual agents in public health contexts. Emphatically, such claims of duties and rights do not solely apply to frontline medical practitioners, but also those involved in using AI in medical research, training large language models on medical data, or advocating and lobbying for greater incorporation of AI across the board.

What would such discretionary duties look like, across each of these areas?

On Beneficence and Non-maleficence, there exists a *duty to ensure group match* – i.e. data models used for medical diagnosis, prognosis, and prescription, should draw upon demographic data that reflects the particularities and sensitivities of their data. As recent work on gendered algorithmic bias enunciates emphatically[214], there must be an active effort to move beyond cis-gender, heterosexual white men as the default of most data models employed in bio-medical research. Indeed, such diversification in data sampling group would also go a long way in shoring up the accuracy and appeal of drug development and therapeutic models in an increasingly diverse world comprising a heterogeneous global consumer class.

On Autonomy, policymakers should legislate to protect what John Tasioulas terms the *right to a human decision.*[215] This amounts to the right to have a decision undertaken by a human, as opposed to an AI system. Indeed, Joseph Weizenbaum has argued that there are certain decisions no computer should be permitted to take.[216] A weaker version of this thesis is that all stakeholders within medical contexts – first and foremost patients – possess *a second order right to opt out of AI treatment, and*

[213] See the distinction between infringement and violation. See Thomson, Judith Jarvis, "Some Ruminations on Rights," *Arizona Law Review* 19, (1977): 45 – 60; Oberdiek, John, "Lost in Moral Space: On the Infringing/Violating Distinction and its Place in the Theory of Rights," *Law & Philosophy* 23, (2004): 325 – 346.
[214] Criado Perez, Caroline, *Invisible Women: Data Bias in a World Designed for Men,* (New York: Abrams Press, 2019).
[238] Tasioulas, John, "Q&A with John Tasioulas," *AI2050 Community Perspectives*, July 14, 2023, https://ai2050.schmidtsciences.org/community-perspective-john-tasioulas/.
[216] Weizenbaum, Joseph, *Computer Power and Human Reason: From Judgment to Calculation,* (San Francisco: W. H. Freeman and Company, 1976).

to opt into human decision. Going forward, medical and public health systems should allow for patients to consent to and opt out of AI usage across disparate stages (junctures) of diagnosis, prognosis, and prescription.

The requirement of patient-centric consent and endorsement of AI deployment in medical treatment can be normalised in the same way as the "Do Not Resuscitate" (DNR) Advance Directives that patients already can opt into – in the event they would not like to have cardiopulmonary resuscitation (CPR) in the event that their hearts stop beating.[217] A "No AI Order" (NAO) may be similarly appropriate and contextually justified.

With regards to Justice, individuals should, in theory, possess access to AI irrespective of their socioeconomic, ethnic, gender backgrounds and other immutable characteristics. Yet it is more difficult in practice to justify the asserted existence of an innate right against the interference by the birth lottery in differential health outcomes – given the claimability objection, as advanced by Onora O'Neill and others.[218] Granting this, whilst such inequalities may be inevitable in practice, efforts must be undertaken to close the gap – especially given the increasingly salient concern that AI may give rise to greater health inequities in the primary care sector[219], and beyond. Justice and equality are not reducible into one another. However, the overt imposition and entrenchment of inequalities – through the amplification of resource-based disparities – is clearly unjust.

Undergirding all these prescriptions, and in ensuring that deliberation and debate can be meaningfully held, individuals must be afforded the *right to minimal comprehension vis-à-vis AI.* Not only do they have the right to know the AI systems involved in their care; they should also be given sufficient supplementary information for them to understand – at least above a minimal threshold – the logic underlying the mechanisms with which AI is incorporated into their healthcare.

[217] MedlinePlus Medical Encyclopedia, "Do-not-resuscitate Order,", reviewed February 3, 2024, https://medlineplus.gov/ency/patientinstructions/000473.htm.
[218] O'Neill, Onora, *Bounds of Justice* (Cambridge: Cambridge University Press, 2000). Also see Stemplowska, Zofia, "Is Humanity under a Duty to Deliver Socioeconomic Human Rights?," *Journal of Applied Philosophy* 39 (2), (2022): 202 – 211 and Estlund, David, *Utopophobia: On the Limits (If Any) of Political Philosophy* (Princeton, NJ: Princeton University Press, 2019).
[219] D'Elia, Alexander, Mark Gabbay, Sarah Rodgers, Ciara Kierans, Elisa Jones, Irum Durrani, Adele Thomas, and Lucy Frith, "Artificial Intelligence and Health Inequities in Primary Care: A Systematic Scoping Review And Framework," *Family Medicine and Community Health* 10, supplementary 1 (2022): e001670.

Part 7: Not Jurisdiction-specific – 2

Ethical Considerations for Trustworthy AI in Healthcare

Ms Tanya Brigden
Senior Policy Analyst (Biomedical Ethics), PHG Foundation, University of Cambridge, United Kingdom

It is anticipated that the use of Artificial Intelligence (AI) could transform healthcare. Both for patients who might benefit from more accurate and effective disease detection, diagnosis, and treatment, and for clinicians who could use AI to support medical decision-making, provide real-time assistance and insights, and reduce time spent on administrative tasks.

AI driven tools are already being used in a healthcare context in the United Kingdom (UK), albeit in a relatively limited way. A recent survey by the Turing Institute reported that over a quarter of the 929 registered doctors surveyed (29%) have used some form of AI in their practice in the last 12 months,[220] with more than half (52%) optimistic about its use in healthcare. Despite this optimism, they also highlighted a number of concerns around using AI systems, including not fully understanding the risks, uncertainty around responsibility for decision making, and a lack of adequate training to understand their professional responsibilities.[221]

Researchers and AI developers have often claimed that "trust" is a critical determinant of the successful adoption of AI in medicine, disputed by philosophers and academics who argue that the emphasis needs to be shifted away from trying to secure trust. Instead, we must direct our efforts toward demonstrating *trustworthiness,* providing the public with the ability to trust intelligently.[222] There is debate around the extent to which AI tools themselves can be considered trustworthy, as being trustworthy is more than just being predictable and reliable; rather, it signifies a moral characteristic or virtue, lacking in inanimate objects such as AI systems.[223] For this reason, it may be more coherent to describe AI systems as "responsible" or "having confidence in" these systems. Nevertheless, the ethical adoption of these tools relies upon the organisations and individuals responsible for developing and deploying AI being trustworthy. This idea, influenced by philosopher Onora O'Neill among others, has been very influential, with the term "trustworthy AI" being widely adopted by the research community, public sector organisations and bodies issuing AI ethics guidance.[224]

There are a number of interconnected ethical questions that are important in fostering trustworthiness in the use of medical AI. To what extent does the AI tool need to be transparent and explainable? How can AI be used equitably in healthcare pathways? What is the nature and level of human involvement that may be appropriate? And, who is responsible for decisions made using AI tools? I shall briefly consider each of these in turn.

---

[220] These were primarily diagnostic and decision support systems, such as image processing and risk assessment, and generative AI tools (most commonly different versions of ChatGPT from OpenAI).
[221] Hashem, Youmna, Saba Esnaashari, Deborah Morgan, John Francis, Anton Poletaev, Florence Enock, and Jonathan Bright, "One in Four UK Doctors Are Using Artificial Intelligence: Exploring Doctors' Perspectives on AI After the Emergence of Large Language Models", *The Alan Turing Institute,* 2024.
[222] O'Neill, Onora, "Linking Trust to Trustworthiness", *International Journal of Philosophical Studies* 26, no.2 (2018): 293–300. https://doi:10.1080/09672559.2018.1454637.
[223] Dlugatch, Rachel, Antoniya Georgieva and Angeliki Kerasidou. "Trustworthy Artificial Intelligence and Ethical Design: Public Perceptions of Trustworthiness of an AI-Based Decision-Support Tool in the Context of Intrapartum Care", *BMC Medical Ethics* 24, no. 42 (2023). https://doi.org/10.1186/s12910-023-00917-w.
[224] European Commission: Directorate-General for Communications Networks, Content and Technology, "Ethics Guidelines for Trustworthy AI", Publications Office (2019). https://data.europa.eu/doi/10.2759/346720.

Explainability — the capacity to express why an AI system reached a particular decision, recommendation, or prediction — is widely considered to be a pillar of trustworthy AI. This is due to the fact that transparent and explainable systems make the "very evidence needed to place or refuse trust intelligently more available."[225] The field of "explainable AI" aims to overcome the "black box problem" and make deep learning more transparent. This can increase confidence in a model by allowing health care professionals to "see" what the model is detecting, enabling a doctor to cross-reference the AI's findings with their own expertise, and ensure that patterns detected by the AI are clinically meaningful.

There are, however, several arguments against a blanket requirement for explainability that are worth considering. First, "the explainability paradox" refers to the challenge that AI models which achieve better performance are generally less explainable due to the opacity of layers of complex neural networks. If explainability is an essential requirement of use in clinical practice, then this may act as a barrier to deploying some effective AI tools. Second, there is a risk of anthropomorphising AI models and assuming that humans can always understand the explanations they provide, which may not necessarily be the case. An example of this is AI's capability to determine biological sex from retinal scans, a trait that is not currently recognised by ophthalmologists. Saliency maps show the highlighted regions that contributed to the model's decision, but not which features are important.[226] Third, it is not clear the extent to which explainable AI has the potential to influence human decision making, with some preliminary research showing that it could worsen automation bias in some circumstances. Given these challenges, it may be that AI tools which pose a higher clinical risk demand a higher degree of explainability. For example, AI systems carrying out administrative tasks may not need to be as explainable as those adopted for cancer detection, so long as they have been proven to be safe and effective. Explainability is undeniably an important safeguard, however the extent to which an AI model needs to be explainable may be context dependent, taking into consideration how it will be used in the clinical pathway, the degree of risk to the patient and evidence of clinical utility.

Prioritising explainability may be desirable in order to detect and address biases. The use of AI in healthcare has the potential to amplify and systematise biases engrained in medical practice, exacerbating health inequalities. In other words, the use of AI could lead to better outcomes for some groups and worse outcomes for others. This has been exemplified by several AI systems that have shown the ability of algorithms to systematically misrepresent and exacerbate health problems in minority groups.[227] Bias can occur throughout the AI lifecycle (from the problem question that AI is addressing, through to algorithm design), and a significant and widely recognised source of bias is the data underpinning AI systems. It is often necessary to train models with large quantities of data, which means datasets are often sourced to prioritize sample size.[228] Therefore, many health datasets do not adequately represent minority groups. For example, databases used to develop or train AI systems for skin cancer diagnosis contain very few images of people with dark skin,[229] and so have

[225] O'Neill, Onora, "Transparency and the Ethics of Communication" in *Transparency: The Key to Better Governance,* eds. Christopher Hood and David Heald, p. 85–86. (London: British Academy Scholarship, 2006).
[226] Delavari, Parsa, Gulcenur Ozturan, Lei Yuan, Özgür Yilmaz, and Ipek Oruc, "Artificial Intelligence, Explainability, and the Scientific Method: A Proof-of-Concept Study on Novel Retinal Biomarker Discovery", *PNAS Nexus* 2, no. 9 (2023): pgad290. https://doi.org/10.1093/pnasnexus/pgad290.
[227] Seyyed-Kalantari, Laleh, Haoran Zhang, Matthew B. A. McDermott, et al., "Underdiagnosis Bias of Artificial Intelligence Algorithms Applied to Chest Radiographs in Under-served Patient Populations", *Nature Medicine* 27 (2021): 2176–2182. https://doi.org/10.1038/s41591-021-01595-0.
[228] Ganapathi, Shaswath, Jo Palmer, Joseph E. Alderman, et al., "Tackling Bias in AI Health Datasets through the STANDING Together Initiative", *Nature Medicine* 28, no. 11 (2022): 2232–2233. https://doi.org/10.1038/s41591-022-01987-w.
[229] Wen, David, Saad M Khan, Antonio Ji Xu, et al., "Characteristics of Publicly Available Skin Cancer Image Datasets: A Systematic Review", *Lancet Digital Health* 4 (1) (2022): e64–74. https://doi.org/10.1016/S2589-7500(21)00252-1.

been shown to have higher rates of false negatives for darker skin types. This problem is compounded by barriers to accessing health services, which mean that underserved groups are left out of "real-world datasets". Additionally, data from underserved populations are more likely to be incomplete or inaccurate.[230]

Even where an AI tool demonstrates high accuracy overall, this may hide poor performance in some groups. Demonstrating trustworthiness necessitates that steps are taken to mitigate against biases, including requirements for transparent and clear reporting of limitations and biases of datasets, and that governance mechanisms that incentivise the curation and use of datasets for AI systems are diverse, inclusive, and promote AI generalisability.

One of the ways in which regulators safeguard against harms arising from the use of AI systems is through human oversight. Not only does keeping a human-in-the-loop act as a safety net against potential errors or oversights, AI systems lack the ability to understand cultural, social, and individual nuances that might affect health outcomes. For these reasons, qualitative research with patients shows that having a human involved tends to foster greater trustworthiness than solely automated decision making.[231] However, solely automated processing is gaining accuracy and reliability and also offers tangible benefits in terms of speed and cost-effectiveness. In the context of an overstretched health system, such as the National Health Service in the UK, it may not always be reasonable and proportionate to see human intervention as a legitimate safeguard, particularly where solely automated pathways outperform human experts. Careful consideration around the role and function of a human-in-the-loop is needed. This could range from humans checking or reviewing suggestions made by AI systems, to treating AI systems as if they are a colleague within a multidisciplinary team,[232] to substituting judgement.

Finally, consideration of liability and accountability will be important for showing evidence of trustworthiness. AI are considered to be decision support tools rather than agents, as current approaches suggest that medical practitioners will largely shoulder responsibility and liability for harm. However, some commentators have expressed concern that clinicians might be at risk of becoming "liability sinks" unfairly absorbing legal liability for errors and adverse outcomes over which they have limited control.[233] This raises the question of how accountability should be assigned as AI automation develops along a clinical pathway, and how responsibility should be shared across all those involved in the design, institution, running, and use of the system.

[230] Gianfrancesco, Milena A., Suzanne Tamang, Jinoos Yazdany, and Gabriela Schmajuk, "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data", *JAMA Internal Medicine* 178, no.11 (2018):1544–1547. https://doi.org/10.1001/jamainternmed.2018.3763.
[231] Thornton, Nell, Ahmed Binesmael, Tim Horton, and Tom Hardie, "AI in Health Care: What Do the Public and NHS Staff Think?" *The Health Foundation.* 31 July 2024. https://www.health.org.uk/reports-and-analysis/analysis/ai-in-health-care-what-do-the-public-and-nhs-staff-think.
[232] Ulfert, Anna-Sophie, Eleni Georganta, Carolina Centeio Jorge, Siddharth Mehrotra, and Myrthe Tielman, "Shaping a Multidisciplinary Understanding of Team Trust in Human-AI Teams: A Theoretical Framework", *European Journal of Work and Organizational Psychology* 33, no.2 (2023):158–71. https://doi.org/10.1080/1359432X.2023.2200172.
[233] Lawton, Tom, Phillip Morgan, Zoe Porter, et al., "Clinicians Risk Becoming 'Liability Sinks' for Artificial Intelligence", *Future Healthcare Journal* 11, no.1 (2024):100007. https://doi.org/10.1016/j.fhj.2024.100007.

In summary, AI could contribute to clinical decision-making in a variety of healthcare contexts. However, those involved in the development and deployment of AI systems must create a trustworthy environment where AI can be harnessed in a way that is safe and maximises its benefits. In order to do this, we need to grapple with a set of ethical and legal questions that hinder the trustworthiness of the systems governing AI, including around identifying and mitigating bias, what it means to have meaningful human involvement and clarity around responsibility and liability.

Part 7: Not Jurisdiction-specific – 3

The AI as Expert – Some Epistemological and Moral Considerations

Dr Pete Mills
Director, PHG Foundation, University of Cambridge, United Kingdom

Artificial Intelligence (AI) systems ("machines") have the potential to make useful contributions to complex clinical cases. There are, however, epistemological, ontological and moral differences between machines and human clinicians, which may have a significant bearing on their interactions.

Contemporary medicine has become an increasingly complex and technically specialised field of human activity. This is not just a result of the increase in the quantity of knowledge and data but also represents a proliferation in ways of knowing: the emergence of new and distinct technical fields, with their own languages and epistemologies. Modern medicine turns the model of "one doctor-many patients" on its head; we may now bring to bear the expertise of many specialists on each patient, with the patient rather than the doctor at the centre. Hence the increasing prevalence, in modern medicine, of multidisciplinary teams (MDTs) that combine expert knowledge from a range of medical specialties and scientific disciplines. This leads to a particular kind of challenge: that of integrating and applying knowledges and forms of expertise from multiple sources in any given clinical case. It also increases the human resource input, since there are now a number of experts involved in assessing, diagnosing, managing, and supporting each patient.[234] At least some of these areas of expertise are now becoming tractable to machine intelligence.[235]

In some areas, notably military applications, significant programmes of research have been carried out over the last decade or more into human-AI teaming.[236] However, while there are some studies of human interaction with health recommender systems, incorporating machines as distinct members of an MDT in clinical medicine has not yet been widely discussed. A recent paper envisages the appearance of "AI thought partners" in the field of medicine, which would "both understand us—reasoning about the doctor, patient and care team as agents with goals, beliefs and worries—and complement our capabilities, integrating swaths of evidence that exceed our cognitive capacities to inform diagnosis and treatment."[237] In the authors' view, beyond considerations such as efficiency, accuracy, robustness, fairness etc., what is distinctive about a machinic thought partner is to be found in its particular way of relating to its human partners. They propose three desiderata for human-AI thought partnerships that may be glossed as:

---

[234] A genetics MDT might include, for example: clinical geneticists, genetic counsellors, molecular and cytogeneticists, bioinformaticians, pathologists, pharmacologists, clinical scientists, disease specialists (e.g. specialists in metabolic diseases), and specialist nurses, as well as ethicists and social workers.

[235] In order to get to this question I have rather obviously vaulted over the currently live question of integrating automated machine processing into distinct disciplines such as digital pathology. This was touched on in the essay by my colleague, Tanya Bridgen.

[236] The US Defense Advanced Research Projects Agency (DARPA), for example, set out to explore ways of effectively integrating human and machine actors into military teams (described as "agile teams" or "A-teams"). Defense Advanced Research Projects Agency, "A-Teams: Agile Teams," Defense Advanced Research Projects Agency, effective 16 December 2024, https://www.darpa.mil/research/programs/agile-teams. For a narrative review of extant empirical research on Human-Autonomy Teams (HATs) see O'Neill, Thomas A., Nathan J. McNeese, Amy Barron and Beau Schelble. "Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature," *Human Factors* 64, no. 5 (2022): 904–38. https://doi.org/10.1177/0018720820960865.

[237] Collins, Katherine M., Ilia Sucholutsky, Umang Bhatt et al., "Building Machines that Learn and Think with People," *Nature Human Behaviour* 8, (2024) 1851–63. https://doi.org/10.1038/s41562-024-01991-9.

(1) that the machine "understands" the human partner, their goals, plans, beliefs, values and circumstances, in effect having a "theory of mind";

(2) that the human partner can "understand" the machine because it communicates in an intuitively understandable way;

(3) that the machine partner is grounded in reality and operates with a true and shared representation of the world.

These desiderata help to illuminate some of the challenges involved in human-AI teaming in general. In the first place, we do not know that they share our model of the world; they are certainly not "tethered" to it in the way that we are.[238] Secondly, for the same reason we may not know that machines have a theory of the world, we are equally unable to know that they have a theory of mind – although they are certainly capable of acting as though they have both. Thirdly, however, we do have something that can plausibly stand for intuitively graspable communication and this is the narrow defile through which our interactions are obliged to pass.

There is, of course, plenty of space for scepticism to enter into any claims we might entertain about machines achieving knowledge and applying it reliably in an interaction with a human interlocutor or with the world more generally. For example, the philosopher Hubert Dreyfus argued that the claims made by AI pioneers for early expert systems were compromised by false assumptions about the nature of thinking and its relationship to lived experience.[239] This critique led to a fruitful line of inquiry about the nature of expertise that was taken forward by the sociologist Harry Collins in the 1990s, self-consciously in dialogue with Dreyfus.[240] What Collins and Dreyfus have in common is a commitment to the idea that the orientation of knowledge towards the world, in a way that is effective in resolving certain types of concrete problems, depends on an unconscious or tacit background of understanding. But whereas Dreyfus insists that expertise had to be grounded in an individual's embodied experience of the world, Collins, working with Rob Evans in Cardiff, developed a theory of expertise that shifts this background to the collective use of language.[241] Collins and Evans allow that a certain kind of "interactional expertise" in a practice language may be acquired without direct experience of that practice. (Interactional expertise is roughly the ability to interact competently with acknowledged experts in a particular field, in a way that would enable the "interactional expert" to pass a relatively demanding Turing Test in that area of expertise. They distinguish this from the more exacting "contributory expertise", which requires sustained individual practice.) Collins does not think embodiment is irrelevant to the acquisition of expertise but he thinks it is less relevant to the individual actor, who is able to access expertise through their socialisation in a living, collective language community, so long as that collective language is fed by practice. Thus, though this practice depends on embodiment, it need not be one's own. The question that remains hanging is that of what kind of minimal embodiment is required to gain a foothold in this language community in the first place.[242]

Having broached the significance of embodiment in relation to knowledge I want quickly to re-trace this line of thought on a different ground, that of ethics. A prominent ethical rubric that covers human-

[238] Collins, Katherine M., Ilia Sucholutsky, Umang Bhatt et al., "Building Machines that Learn and Think with People," *Nature Human Behaviour* 8, (2024) 1851–63. https://doi.org/10.1038/s41562-024-01991-9.
[239] Dreyfus, Hubert L., *What Computers Still Can't Do, A Critique of Artificial Reason* (Cambridge: The MIT Press, 1992).
[240] Collins, Harry, "Interactional Expertise and Embodiment," in *Skillful Performance: Enacting Capabilities, Knowledge, Competence, and Expertise in Organizations,* ed. Jorgen Sandberg, Linda Rouleau, Ann Langley, and Haridimos Tsoukas (Oxford University Press, 2017).
[241] Collins, Harry and Robert Evans, *Rethinking Expertise* (University of Chicago Press, 2007).
[242] See Collins, Harry. "Language and Practice." *Social Studies of Science* 41 no.2 (2011) 271–300. https://doi.org/10.1177/0306312711399665.

autonomy teaming is that of "human-centred AI". Human-centred AI systems are designed to work with and for people, in complex, contextually rich and ambiguous, "high-stakes" real world situations.[243] This seems particularly relevant to the practice of medicine, and has very clear resonances with the idea of patient-centred care, with which we began. A prominent theme in this area is that of trust between human and machine team members. However, the idea of "trust" in this context seems a jarringly – perhaps deliberately – anthropic concept to use in this connection, at least insofar as one accepts that it has an inherent moral component. Trusting is arguably different from "relying on" or "having confidence in", comprising, alongside performative ability, qualities such as integrity and benevolence.[244] This gives the concept of trust a kind of moral thickness that, I would argue, applies only among members of a moral community.[245] Given more space, I would argue that embodiment in the world, in a body that is minimally similar to others' in the community, is a necessary condition of belonging to this community.

What is the significance of this for the incorporation of AI into multidisciplinary clinical teams? I certainly do not mean to deny that a reliable autonomous agent has considerable potential to contribute to medical practice. But if we can integrate machine intelligence into teams at all, we can theoretically turn the problem inside out and ask why, if successive areas of expertise fall to be rendered by machines, machine intelligence could not integrate all medical practice. I think the answer is to be found in the kind of practice that medicine is. There must be a human in the loop, not merely to govern but to ground that practice as a clinical and moral one.[246] How stretched or attenuated this loop may be, however, is a question that merits further attention.

[243] See, for example, Barmer, Hollen, Rachel Dzombak, Matt Gaston, Jay Palat, Frank Redner, Carol Smith and Tanisha Smith, *Human-Centered AI*, (Carnegie Mellon University Report, 2021). https://doi.org/10.1184/R1/16560183.v1.
[244] See, for example, McNeese, Nathan J., Mustafa Demir, Erin Chiou, Nancy Cooke and Giovanni Yanikian. "Understanding the Role of Trust in Human-Autonomy Teaming." *Proceedings of the 52nd Hawaii International Conference on System Sciences* (2019): 254–63. https://hdl.handle.net/10125/59466.
[245] Compare, for example, the statements: (A) "I am confident that you will betray me" and (B) "I trust you to betray me." In English, at least, the second involves a kind of paradox, since the concept of trust implies a sympathy that is shared by moral agents.
[246] There is one final and intriguing possibility – that that human in the loop could be the patient themselves. This is a question that certainly seems to be worthwhile exploring but not one that there is space to explore here.

Part 7: Not Jurisdiction-specific – 4

Emerging Ethical Challenges in Medical AI: Persuasion, Manipulation, and Consent

Prof Rachel Sterken
Associate Professor, Department of Philosophy
Co-director of ConceptLab
The University of Hong Kong, Hong Kong SAR, China

The integration of Artificial Intelligence (AI) into healthcare has ushered in transformative advancements, particularly through AI-driven diagnostic tools and patient care systems. However, this technological change presents risks and ethical challenges, notably concerning persuasion, manipulation, and informed consent. As AI technologies integrate into healthcare, these risks necessitate scrutiny, planning, and governance. This report examines these issues, emphasizing their implications for medical professionals and patients.

AI applications in healthcare span a variety of fields. From radiology (e.g., algorithms demonstrate exceptional proficiency in analyzing medical images, sometimes outperforming human specialists),[247] to cardiology (e.g., AI systems detect arrhythmias in electrocardiograms with remarkable precision),[248] to dermatology (e.g., AI systems excel at classifying melanoma in dermoscopic images),[249] these advancements highlight the transformative potential of AI in healthcare. However, they also underscore the need to balance technological progress with ethical considerations, particularly as these systems begin to influence the core principles of medical practice. For the remainder of this report, the term AIMD shall be used to refer to "Medical Devices with AI or Machine Learning Capabilities" that are employed in healthcare contexts.

Trust emerges as a central theme when considering the integration of AI into healthcare. Medical professionals may struggle to trust AIMDs due to their opaque decision-making processes. This lack of transparency, often described as a "black-box" phenomenon, leaves clinicians questioning how and why certain conclusions are reached. Similarly, patients face challenges in trusting results or outputs of AIMDs, or even of medical professionals who utilize AIMDs. These dynamics highlight the dual-layered complexity of trust in AI-mediated healthcare. This erosion in trust arises due to numerous factors, including the persuasive and manipulative capabilities that AIMDs may exhibit.

Persuasion is here understood as a communicative act aimed at influencing the attitudes and action of others. It goes beyond merely conveying information and seeks to shape how information is received and acted upon. The kinds of influences that are persuasive span the conscious/unconscious, rational/nonrational, verbal/nonverbal. On this level, persuasive technologies (PTs) represent an ethical challenge in healthcare. These systems, designed to influence behaviours like exercise or medication adherence, are increasingly pervasive in both medical and consumer contexts. While persuasion is less ethically fraught than manipulation, it still raises concerns about autonomy and

[270] Chan, Heang-Ping et al., "Deep Learning in Medical Image Analysis", in *Deep Learning in Medical Image Analysis*, edited by Gobert Lee and Hiroshi Fujita, vol. 1213, Advances in Experimental Medicine and Biology (Cham: Springer International Publishing, 2020), 3–21, https://doi.org/10.1007/978-3-030-33128-3_1.
[271] Hannun, Awni Y. et al., "Cardiologist-Level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network", *Nature Medicine* 25, no. 1 (January 2019): 65–69, https://doi.org/10.1038/s41591-018-0268-3.
[272] Brinker, Titus J. et al., "Deep Learning Outperformed 136 of 157 Dermatologists in a Head-to-Head Dermoscopic Melanoma Image Classification Task", *European Journal of Cancer* 113 (May 2019): 47–54, https://doi.org/10.1016/j.ejca.2019.04.001.

freedom of thought. The "attention economy,"[250] which monetizes human focus, illustrates how PTs can infringe on personal agency, even in well-intentioned applications.

When persuasion has detrimental outcomes to those being persuaded, or is morally questionable, it is generally labelled as manipulation.[251] Manipulation is ethically problematic since it infringes on a subject's mental sovereignty, undermines informed consent, and often involves covert actions that exploit individual vulnerabilities. These ethical violations pose serious risks in healthcare, particularly when AI systems are designed to subtly influence decisions, operate without transparency, or are given undue authority. The potential for AI systems to engage in manipulation extends beyond intentional design. AI tools can be explicitly programmed for manipulation, such as gamification techniques that exploit behavioral tendencies. More concerning is the phenomenon of algorithmic manipulation, where systems autonomously manipulate users without direct human intervention. This raises complex questions about accountability and ethical boundaries, particularly in contexts where manipulation may harm medical professionals or patients.

A specific mention should here be made to "AI Doctors" with systems such as those based on Generative AI (GenAI) leveraging Large Language Models (LLMs) like ChatGPT. While these platforms can provide patient-facing consultations and assist medical professionals with diagnostic recommendations, their capabilities are not without risks. Known issues with LLMs include their propensity to hallucinate, offering fabricated information; their sycophantic tendencies, reinforcing biases or assumptions from users; and their detachment from real-world context, which can lead to outdated or irrelevant medical advice. Perhaps most concerning is the potential for these systems to deliver diagnoses without physical examination, raising ethical questions about the validity and reliability of such practices.

Given the above, therefore, three issues of concern can be identified which should be mitigated against:

Transparency: Transparency is one of the cornerstones of ethical AI deployment, yet it remains elusive in many AIMDs. These systems often produce accurate analyses or diagnoses, but without a clear explanation of their reasoning. The lack of transparency creates what is referred to as "black-box medicine,"[252] "in which the basis for a given output is not always sufficiently clear and thus complicates its evaluation in view of potential errors and biases of the system, arising, for example, from the quality and breadth of data it has been trained with."[253] This opacity not only challenges medical professionals' ability to trust and rely on AI systems but also undermines efforts to establish robust governance structures.

Responsibility: When AIMDs are involved in medical care, another issue arises pertaining to who is ultimately responsible for the patient's care. More importantly, determining who is ultimately responsible for errors—whether the clinician, the device, or its manufacturer—becomes a contentious and unresolved question. Furthermore, the opacity mentioned above also exacerbates issues when

[273] Lanham, Richard A., *The Economics of Attention: Style and Substance in the Age of Information* (Chicago: University of Chicago Press, 2007).
[274] Nettel, Ana Laura and Georges Roque, "Persuasive Argumentation Versus Manipulation", *Argumentation* 26, no. 1 (2012): 55–69, https://doi.org/10.1007/s10503-011-9241-8; Noggle, Robert, "The Ethics of Manipulation", in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (Stanford: Metaphysics Research Lab, Stanford University, 2022), https://plato.stanford.edu/archives/sum2022/entries/ethics-manipulation/.
[275] Price II, W. Nicholson, "Black-Box Medicine", *Harvard Journal of Law and Technology* 28, no. 2 (2015): 420-467.
[276] Braun, Matthias et al., "Primer on an Ethics of AI-Based Decision Support Systems in the Clinic", *Journal of Medical Ethics* 47, no. 12 (December 2021): e3, https://doi.org/10.1136/medethics-2019-105860.

medical professionals and AIMDs might disagree.[254] This epistemic worry highlights another possibility of manipulation. AIMDs, in the way outputs are given and perceived, may unduly influence medical professionals into rejecting their own diagnosis and accepting that of AIMDs.

Patient Autonomy and Informed Consent: The use of AIMDs may also potentially threaten patient autonomy, particularly when the recommendations of AIMDs overshadow patients' personal values. One must be wary of a "computer-knows-best"[255] mentality, where machines dictate treatment options based on their programmed priorities rather than the patient's individual preferences. This dynamic exemplifies a form of manipulation that compromises the fundamental principles of patient-centred care. In a related vein, a critical aspect of patient autonomy is informed consent. Patients may be unaware that their diagnoses or treatment plans are fully or partially the result of AI systems, and whether or not clinicians are morally obliged to disclose this fact to patients is not entirely clear.[256]

Thus, AIMDs, along with other elements in Medical AI more widely, present ethical challenges, as exemplified in the specific areas of persuasion, manipulation, and consent explored above. The potential benefits of these technologies are transformative, and to refrain from using them would be a disservice to medical practice and patients. However, these benefits can only be reaped if AIMDs are implemented and used in a responsible and safe manner. Thus, the examination and mitigation of the ethical implications that arise from the use of such tools is of great importance. Addressing these challenges requires a twofold approach:

First, a concerted effort must be made to *expand research into the ethical impacts of AIMDs*, providing a robust foundation for understanding and mitigating risks. Given the rapid pace of advancement in this field (as in other AI-related fields more generally), the need for interdisciplinary research in the various stages of the implementation of AIMDs – from developers, deployers and users – is all the more urgent.

Second, *legislative frameworks as well as procedural "best-practices" must evolve* to provide clearer guidance and accountability mechanisms. Legislation such as the European Union's AI Act represents important steps forward but remains ambiguous about the specific risks that warrant scrutiny.[257] A more precise and proactive approach to regulation, informed by interdisciplinary research and stakeholder engagement, is essential to ensure that AIMDs are developed and deployed in a manner that safeguards trust, autonomy, and equity in healthcare.

---

[277] Kempt, Hendrik and Saskia K Nagel, "Responsibility, Second Opinions and Peer-Disagreement: Ethical and Epistemological Challenges of Using AI in Clinical Diagnostic Contexts", *Journal of Medical Ethics* 48, no. 4 (April 2022): 222–29, https://doi.org/10.1136/medethics-2021-107440.
[278] McDougall, Rosalind J, "Computer Knows Best? The Need for Value-Flexibility in Medical AI", *Journal of Medical Ethics* 45, no. 3 (March 2019): 156–60, https://doi.org/10.1136/medethics-2018-105118.
[279] Hatherley, Joshua, "Are Clinicians Ethically Obligated to Disclose Their Use of Medical Machine Learning Systems to Patients? ", *Journal of Medical Ethics*, 7 August 2024, jme-2024-109905, https://doi.org/10.1136/jme-2024-109905.
[280] Aboy, Mateo, Timo Minssen, and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7, no. 1 (6 September 2024): 237, https://doi.org/10.1038/s41746-024-01232-3.

Analysis & The Way Forward

The following is (a) an analysis of the expanded abstracts, presentation transcript and presentation summary in this report and discussions at the conference and (b) a potential way forward for the governance of medical AI in Hong Kong.

1. Introduction

We are at a critical juncture in the global effort to regulate and govern the use of medical Artificial Intelligence (AI). As AI technologies demonstrate increasing promise across a range of healthcare applications — from diagnostics and treatment recommendations to administrative efficiencies — the need for robust, context-sensitive governance frameworks has never been more urgent. In Hong Kong, there is currently no single, comprehensive legislation governing medical AI. Hong Kong can learn from and adapt international regulatory models to shape a forward-looking approach that balances innovation, patient safety, and public trust in overcoming the challenge of governing medical AI.

The approaches currently adopted in the different jurisdictions that feature in this report — Hong Kong SAR, Macau SAR, Mainland China, the European Union (EU), the United Kingdom (UK), Canada, the United States (US), Qatar, Saudi Arabia, and the United Arab Emirates — are at various stages of developments, with some more committed to one path over another. Most of these jurisdictions have so far avoided introducing comprehensive legislation dedicated to the regulation of AI. This section provides an analysis of medical AI regulation and governance in these jurisdictions, including key developments, benefits and drawbacks as raised by the conference speakers.

2. The current approach in Hong Kong SAR

The Government of the Hong Kong Special Administrative Region of the People's Republic of China (Government of the HKSAR) has adopted a multi-pronged approach to develop an AI ecosystem. This approach was outlined by Mr Donald Mak, Deputy Commissioner for Digital Policy (Data Governance) at the Digital Policy Office (DPO) of the Government of the HKSAR, who highlighted the Government's "Ethical Artificial Intelligence Framework"[258] as a reference for guiding AI adoption and the Government's "Hong Kong Generative Artificial Intelligence Technical and Application Guideline"[259] ("DPO Guideline"). The "Ethical Artificial Intelligence Framework" establishes foundational principles, such as transparency and interpretability, accountability, and fairness. As described in its preamble, the DPO Guideline "documents the technical background and governance principles of generative AI, and provides a practical guide for Technology Developers, Service Providers and Service Users". It should be noted that the DPO Guideline has set out recommendations for, among others, the healthcare industry.[260] As pointed out by Mr Donald Mak, the DPO Guideline features a four-tiered risk classification, with a proportionate management approach based on potential harm and the four tiers of risks being "unacceptable risk", "high risk",

[258] Digital Policy Office of the Government of the Hong Kong Special Administrative Region of the People's Republic of China ("Government of the HKSAR"), "Ethical Artificial Intelligence Framework (Customised Version for General Reference by Public) Version: 1.4", updated July 2024, https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/.

[259] Digital Policy Office of the Government of the HKSAR, April 2025, "Hong Kong Generative Artificial Intelligence Technical and Application Guideline", https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/doc/HK_Generative_AI_Technical_and_Application_Guideline_en.pdf.

[260] *Ibid*., p.40.

"limited risk" and "low risk".[261] Under the "proposed AI governance framework" set out in the DPO Guideline, the regulatory strategies for each of the tiers are as follows:[262]

- "Unacceptable risk": full prohibition; legal liability for development/deployment
- "High risk" (e.g. deployed in the critical infrastructure context of healthcare diagnostics): conformity assessments, human-in-the-loop requirements, real-time monitoring
- "Limited risk": transparency obligations, user opt-out mechanisms, annual compliance audits
- "Low risk": self-certification

The Office of the Privacy Commissioner for Personal Data (PCPD) in Hong Kong has also contributed to this governance landscape through the release of the "Guidance on the Ethical Development and Use of Artificial Intelligence"[263], the "Artificial Intelligence: Model Personal Data Protection Framework"[264] and the "Checklist on Guidelines for the Use of Generative AI by Employees"[265]. These documents provide guidance to assist in ensuring compliance with the Personal Data (Privacy) Ordinance (PDPO) (Cap. 486).

Governance of medical devices is primarily overseen by the Medical Device Division (MDD) within the Department of Health in Hong Kong. Mr Lam Kam Chun, Tommy, Senior Electronics Engineer (Medical Device) in the MDD, pointed out that Hong Kong currently has no overarching legislation governing the manufacture, import, distribution, supply, and use of medical devices, although, depending on their nature and characteristics, some may be regulated by existing legislation. The Medical Device Administrative Control System (MDACS) was introduced by the MDD in 2004. MDACS is a voluntary scheme for the listing of medical devices and traders.[266] Mr Lam explained that this framework combines a premarket listing approach with post-market controls. He added that the MDACS is regularly reviewed, updated and aligned with international standards, drawing on recommendations from the International Medical Device Regulators Forum (IMDRF) and the Global Harmonization Working Party (GHWP). Mr Lam drew attention to a technical guidance published in 2024 by the MDD on the standards set by the MDACS for AI medical devices, Technical Reference TR-008.[267] This reference document states that the classification of a medical device integrated with AI in the form of software (AI-MD) shall follow the risk-based classification principle in accordance with two other technical references and establishes requirements that include, *inter alia*, safety mechanisms and software version control[268].

It should be noted that the Intellectual Property Department (IPD) of the Government of the HKSAR

---

[261] *Ibid*., p.11.
[262] *Ibid*., p.11.
[263] Office of the Privacy Commissioner for Personal Data, Hong Kong, "Home: Resources Centre: Publications: Guidance Notes/ Reports", Guidance on the Ethical Development and Use of Artificial Intelligence, August 2021, https://www.pcpd.org.hk/english/resources_centre/publications/guidance/guidance.html?year=2021.
[264] Office of the Privacy Commissioner for Personal Data, Hong Kong, "Home: Resources Centre: Publications: Guidance Notes/ Reports", Artificial Intelligence: Model Personal Data Protection Framework, June 2024, https://www.pcpd.org.hk/english/resources_centre/publications/guidance/guidance.html?year=2024.
[265] Office of the Privacy Commissioner for Personal Data, Hong Kong, "Privacy Commissioner's Office Publishes (1) Checklist on Guidelines for the Use of Generative AI by Employees and (2) Investigation Findings on the Data Breach Incident of ImagineX Management Company Limited", 31 March 2025, https://www.pcpd.org.hk/english/news_events/media_statements/press_20250331.html.
[266] Medical Device Division of the Department of Health of the Government of the HKSAR, "Medical Device Administrative Control System", revised 4 July 2024, https://www.mdd.gov.hk/en/mdacs/index.html.
[267] Medical Device Division of the Department of Health of the Government of the HKSAR, "Artificial Intelligence Medical Devices (AI-MD) Technical Reference: TR-008", revised on 15 November 2024, https://www.mdd.gov.hk/filemanager/common/mdacs/TR008.pdf.
[268] *Ibid*., p.5 and 7. At present, "Artificial Intelligence Medical Devices (AI-MD) Technical Reference: TR-008" issued by the Department of Health of the Government of the HKSAR is not compulsory for any Artificial Intelligence Medical Device that is not listed in the Medical Device Administrative Control System (MDACS).

conducted a public consultation in 2024 to seek public views on deepfakes, the transparency of AI systems and certain copyright issues relating to AI.[269] In a paper issued in 2025 on the outcomes of the consultation and the proposed way forward, IPD indicated that:

- IPD did not consider it appropriate to address the issue of deepfakes solely from the perspective of copyright or intellectual property (IP), since this issue is "interconnected with a broad range of issues in multiple fields, and involves the application of existing laws under different domains".[270] "For other responses and suggestions relating to the overall transparency of AI systems which falls outside the domain of copyright … given that they cover multiple domains and are not confined to or even go beyond the realm of copyright and IP protection", IPD did not consider it "appropriate to address them separately and solely from the perspective of copyright or IP".[271] The Government of the HKSAR would continue to have a close watch on "the latest development and the international trend, with a view to reviewing and updating the relevant legislation and guidelines". [272]
- IPD considered it necessary to amend the Copyright Ordinance (Cap. 528) to introduce a specific "text and data mining exception" to "allow reasonable use of copyright works for computational data analysis and processing". [273]
- IPD considered that there was "no sufficient justification to propose any legislative amendments concerning the copyright protection and infringement issues relating to AI-generated works" and it planned to formulate guidelines on copyright issues in this regard.[274]

3. Contrasting comprehensive cross-sector regulation of AI against a pro-innovation approach

The conference highlighted that there are two well-characterised diverging approaches to the governance of medical AI. The first is an approach of comprehensive cross-sector regulation, such as that adopted by the EU. The second is a pro-innovation approach, which is more flexible and light-touch, such as that adopted by the UK and the US. Speakers outlined the relevant facets of these approaches, and spoke to their various strengths and weaknesses.

3.1 The EU's comprehensive AI legislation

The EU has been the first to approve comprehensive legislation on AI — the EU Artificial Intelligence Act (EU AI Act). Despite being branded as an "Act", it is an EU Regulation (Regulation EU 2024/1689), and as such it has legal effect in all EU member states without the need for separate national legislation. Prof Timo Minssen, Professor of Law at the University of Copenhagen in Denmark, described the AI Act as representing "a significant effort to regulate AI technologies … balancing the need for safety and efficacy with the imperative to foster innovation". At the same time, Prof Minssen warned that "the potential challenges posed by overlapping regulatory requirements, stakeholder collaboration, and the rapid pace of technological change cannot be overlooked".

---

[269] Intellectual Property Department of the Commerce and Economic Development Bureau of the Government of the HKSAR, "Enhancement of the Copyright Ordinance regarding Protection for Artificial Intelligence Technology Development", LC Paper No. CB(1)999/2024(05), July 2024, https://www.legco.gov.hk/yr2024/english/panels/ci/papers/ci20240716cb1-999-5-e.pdf.

[270] Intellectual Property Department of the Commerce and Economic Development Bureau of the Government of the HKSAR, "Enhancement of the Copyright Ordinance regarding Protection for Artificial Intelligence Technology Development – Outcomes of Public Consultation and Proposed Way Forward", LC Paper No. CB(2)240/2025(04), February 2025, p. 9, https://www.legco.gov.hk/yr2025/english/panels/ci/papers/ci20250218cb2-240-4-e.pdf.

[271] *Ibid*., p.9.

[272] *Ibid*., p.9.

[273] *Ibid*., p.11.

[274] *Ibid*., p.10.

In their presentations, Prof Vera Lúcia Raposo, Associate Professor of Law and Technology at the NOVA School of Law in Portugal, and Prof Timo Minssen outlined the key features of the EU AI Act, and its implications for medical AI.

At the core of the EU AI Act is a risk-based approach to AI regulation, classifying general-purpose AI models into (a) general-purpose AI models and (b) general-purpose AI models with systemic risk and classifying AI systems into the following four categories[275]:

- Prohibited practices in relation to AI systems (e.g. the placing on the market, the putting into service for this purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except for medical or safety reasons)
- High-risk AI systems (e.g. AI systems of certain medical devices, certain AI systems intended to be used to dispatch emergency first response services, including by police, firefighters and medical aid, as well as of emergency healthcare patient triage systems)
- AI systems with specific transparency risk, which are sometimes described as limited risk AI systems
- Minimal risk AI systems

The EU Act expressly excludes certain AI systems, certain AI models and certain entities from its scope of application. For example, Article 2(6) of the EU AI Act provides that this Regulation does not apply to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development. Another example is Article 2(8) of the EU AI Act, which provides that, *inter alia*, this Regulation does not apply to any research, testing or development activity regarding AI systems or AI models prior to their being placed on the market or put into service and that testing in real world conditions shall not be covered by that exclusion.

The EU AI Act imposes differing obligations on various actors, such as providers, deployers, importers and distributors.

The EU AI Act imposes obligations in respect of (a) **general-purpose AI models** and (b) **general-purpose AI models with systemic risk** as well as general-purpose AI systems.[276]

Article 5 of the EU AI Act **prohibits certain AI practices**, including but not limited to the following:

- "the placing on the market, the putting into service for this … purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons"
- "the placing on the market, the putting into service or the use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm"

[275] European Parliamentary Research Service, "Briefing: Artificial Intelligence Act", September 2024, p.3 and 8-10, https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf.
[276] Please see, for example, Chapter V, Recital (85) and Article 75 of the Artificial Intelligence Act of the European Union.

- "the placing on the market, the putting into service or the use of an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm"

Recital (29) of the EU AI Act states, *inter alia*, that "[t]he prohibitions of manipulative and exploitative practices in this Regulation should not affect lawful practices in the context of medical treatment such as psychological treatment of a mental disease or physical rehabilitation, when those practices are carried out in accordance with the applicable law and medical standards".

"**High-risk AI system**" is the highest risk classification for permitted uses of AI systems and classification as a "high-risk AI system" triggers a cascade of regulatory requirements under the EU AI Act, such as those relating to risk management (Article 9), data and data governance (Article 10), transparency and provision of information to deployers (Article 13), human oversight (Article 14), accuracy, robustness and cybersecurity (Article 15), conformity assessment (Article 43) and post-market monitoring (Article 72).

Article 50(1)-(4) of the EU AI Act imposes information and transparency obligations in respect of **"certain AI systems"**, which are sometimes described as **AI systems** posing "**limited risk**".[277] Article 50(6) of the EU AI Act stipulates that Article 50(1)-(4) of the EU AI Act shall not affect the requirements and obligations set out in Chapter III (which is titled "High-risk AI Systems") of the EU AI Act. Hence, the fact that one or more obligations are imposed in respect of an AI system by Article 50(1)-(4) of the EU AI Act does not prevent the obligations imposed in respect of "high-risk AI systems" set out in Chapter III from applying to that AI system.[278] Under Article 50(1) of the EU AI Act, subject to exceptions, providers are required to ensure that AI systems intended to interact directly with natural persons are designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system. Under Article 50(2) of the EU AI Act, subject to exceptions, providers of AI systems, including general-purpose AI systems, generating synthetic audio, image, video or text content, are required to ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated as more particularly described therein. Article 50(3) of the EU AI Act imposes obligations on deployers of an emotion recognition system or a biometric categorisation system, whereas Article 50(4) of the EU AI Act imposes obligations on deployers of an AI system that generates or manipulates image, audio or video content constituting a deep fake and deployers of an AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest.

The EU AI Act is often said to impose no obligations in respect of **"minimal risk" AI systems**[279]. It should be noted that Article 95 of the EU AI Act refers to the drawing up of (a) codes of conduct intended to foster the voluntary application to AI systems, other than high-risk AI systems, of some or all of the requirements set out in Chapter III, Section 2, EU AI Act and (b) codes of conduct concerning the voluntary application "of specific requirements to all AI systems, on the basis of clear objectives and key performance indicators to measure the achievement of those objectives, including

[277] European Parliamentary Research Service, "Briefing: Artificial Intelligence Act", September 2024, p.3 and 8-10, https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf.
[278] Taylor Wessing, "AI Act: High-risk AI Systems – What Applies, What is Due When?", 12 November 2024, https://www.taylorwessing.com/zh-hant/insights-and-events/insights/2024/11/high-risk-ai-systems#:~:text=According%20to%20Article%2050%20(6,for%20high%2Drisk%20AI%20systems.
[279] European Commission, "European Artificial Intelligence Act Comes into Force", 1 August 2024 https://ec.europa.eu/commission/presscorner/detail/en/ip_24_4123.

elements such as, but not limited to ... promoting AI literacy, in particular that of persons dealing with the development, operation and use of AI".

In this connection, it should also be noted that Article 4 of the EU AI Act requires providers and deployers of AI systems to take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf.

The obligations imposed by the EU AI Act in respect of any particular AI system to be used in the medical context, which might or might not be a medical device or constitute part of a medical device, vary depending on which of the aforementioned categories of AI systems it falls into (e.g. whether the system is classified as a "high-risk AI system" under Article 6 of the EU AI Act).

Article 6(2) of the EU AI Act provides that AI systems referred to in Annex III to the EU AI Act shall be considered to be "high-risk", but Article 6(3) of the EU AI Act provides that an AI system referred to in Annex III shall not be considered to be "high-risk" where the requirements set out in Article 6(3) are met.

The following are some of the AI systems currently listed in Annex III to the EU AI Act:

- "AI systems intended to evaluate and classify emergency calls by natural persons or to be used to dispatch, or to establish priority in the dispatching of, emergency first response services, including by police, firefighters and medical aid, as well as of emergency healthcare patient triage systems"
- "AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services"
- "AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance"
- "AI systems intended to be used by or on behalf of competent public authorities or by Union institutions, bodies, offices or agencies to assess … a health risk … posed by a natural person who intends to enter or who has entered into the territory of a Member State" in the context of migration, asylum and border control management, in so far as their use is permitted under relevant Union or national law

Alternatively, an AI system is considered to be a "high-risk AI system" if both of the following requirements are met (Article 6(1) EU AI Act):

(a) the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation listed in Annex I to the EU AI Act; and
(b) the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is mandated to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonisation legislation listed in Annex I to the EU AI Act.

Applying Article 6(1) of the EU AI Act and given that the EU Medical Devices Regulation (EU MDR) and the EU In Vitro Diagnostic Medical Devices Regulation (EU IVDR) are two of the legislation referred to in Annex I to the EU AI Act, an AI system will be classified as a "high-risk AI system" under the EU AI Act by virtue of Article 6(1) of the EU AI Act, if (a) that AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the EU MDR or the EU IVDR and (b) the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is mandated to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the EU MDR or the EU IVDR.

Complex interaction with other regulations

AI medical devices must comply in full with any applicable provisions of both the EU AI Act and, as the case may be, the EU MDR or the EU IVDR, as well as other relevant regulations such as the EU's comprehensive data protection legislation (the General Data Protection Regulation (GDPR)). Speakers emphasised that these layered obligations have the potential to cause confusion and increased regulatory burdens amongst developers, etc. A concern raised by Prof Minssen is that the addition of further regulatory obligations might stifle innovation and create barriers, particularly for small and medium-sized enterprises with limited resources.

Data governance is particularly relevant in the medical field, and the EU AI Act builds on the GDPR to safeguard the data of patients, etc., while introducing AI-specific requirements for training, validation, and testing datasets. For high-risk AI systems, Article 10 of the EU AI Act, *inter alia*, requires training, validation and testing data sets to be subject to data governance and management practices and requires training, validation and testing data sets to be relevant, sufficiently representative, and, to the best extent possible, free of errors and complete in view of the intended purpose. These requirements are important, as the quality of the output of AI systems depends significantly on the quality of such data.

The European Health Data Space

Dr Colin Mitchell, Head of Humanities at the PHG Foundation, described how the European Health Data Space (EHDS) aims to facilitate this through allowing data to be made available and exchanged in a trustworthy and secure manner across the EU. As described by Dr Mitchell, the European Health Data Space Regulation, which entered into force on 26 March 2025 with implementation taking place in phases, sets out the infrastructure, governance and standards which shall apply in the context of electronic health data. This applies to two uses:

- primary use, which means "the processing of electronic health data for the provision of healthcare, in order to assess, maintain or restore the state of health of the natural person to whom those data relate, including the prescription, dispensation and provision of medicinal products and medical devices, as well as for relevant social, administrative or reimbursement services" (Article 2(2))
- secondary use, which means the processing of electronic health data for the purposes set out in Chapter IV of the European Health Data Space Regulation (such as research, innovation, policy-making, and the public interest in the areas of public or occupational health as more particularly described in Chapter IV), other than the initial purposes for which they were collected or produced (Article 2(2))

The EHDS does, however, raise challenges for those seeking to implement it and those established in "third countries" who may wish to interact with it. These include the system of mandatory data disclosure for approved secondary purposes (with a provision stipulating that natural persons shall have the right to opt out and that any member state may provide in its national law for a mechanism to make data for which a right to opt out has been exercised available, provided that the relevant conditions are fulfilled), and the impact this might have on public trust and support; the immense resources needed to develop capacity and digital infrastructure across the EU; and challenges around access to the EHDS by "third country" innovators and researchers. Given these obstacles, it is unclear whether the EHDS can achieve its far-reaching ambitions.

Challenges of a comprehensive AI regulation approach

While the EU AI Act seeks to promote safe and responsible AI technologies, it also raises challenges. These include the need for AI medical devices to comply with applicable regulatory requirements of both the EU AI Act *and*, as the case may be, the EU MDR or the EU IVDR, should the pre-conditions for applicability be met.

Speakers raised the risk of bottlenecks and lack of synchronicity, as the successful implementation of the EU AI Act requires concerted efforts from a broad spectrum of stakeholders, including policymakers, regulators, notified bodies (the independent organisations designated to assess conformity), AI providers, AI deployers and other relevant parties in the supply chain, industry, and the public and private sectors. Prof Minssen illustrated this concern by highlighting that the implementation of the EU MDR has been hampered by a shortage of notified bodies, leading to compliance delays and legal uncertainties. As the AI systems of many commercial AI-enabled medical devices are likely to be classified as high-risk AI systems[280], the increased demand for notified body assessments could overwhelm existing capacities.

Last, there is concern that the dynamic pace of AI innovation might suffer from a static regulatory approach. A number of amendments were required to be made to the proposed EU AI Act in the period between the initial proposal of and the enactment of the EU AI Act, in order to keep pace with the advancement of AI occurring in real time. The initial version of the proposed EU AI Act was regarded as inadequate in terms of governing generative AI, which surged in popularity after the initial proposal of the EU AI Act was put forward. After the initial proposal of the EU AI Act was put forward and before the enactment of the EU AI Act, additional provisions were included in the proposed EU AI Act in order to address generative AI, illustrating the challenges of keeping legislation aligned with technological advancements. As large multi-modal models (LMMs), which are one type of generative AI, rose in recent years, the World Health Organisation issued a guidance in 2024 to help member states in mapping the challenges and benefits associated with the use of LMMs for health and in developing policies and practices for appropriate development, provision and use.[281] This underscores the importance for regulatory frameworks to be sufficiently flexible so that they are capable of accommodating new AI developments and remaining aligned with technological advancements, societal expectations, and ethical considerations at the same time.[282] It could be argued that given how quickly AI technology is advancing, putting in place a comprehensive AI legislation, which is relatively time-consuming to put in place and rigid, at this stage is impractical. It may be that a targeted and progressive regulatory approach[283] which relies more heavily on guidance and existing legislation and puts in place new legal rules only for those pressing issues for which legislation is indispensable at this stage would be better for innovation and competition, fostering an environment

[280] Busch, Felix, Jakob Nikolas Kather, Christian Johner *et al*., "Navigating the European Union Artificial Intelligence Act for Healthcare", *npj Digital Medicine*. 7, 210 (2024). https://doi.org/10.1038/s41746-024-01213-6.
[281] World Health Organization, "Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multi-Modal Models", 2024, p.viii, https://www.who.int/publications/i/item/9789240084759.
[282] Aboy, Mateo, Timo Minssen and Effy Vayena, "Navigating the EU AI Act: Implications for Regulated Digital Medical Products", *npj Digital Medicine* 7, 237 (2024). https://doi.org/10.1038/s41746-024-01232-3.
[283] Research Office of the Legislative Council Secretariat of the Government of the HKSAR, "Information Note: Regulation of Artificial Intelligence in the European Union and the Mainland", 17 February 2025, p.13, https://app7.legco.gov.hk/rpdb/en/uploads/2025/IN/IN04_2025_20250217_en.pdf; Pernot-Leplay, Emmanuel, "The AI Dilemma: AI Regulation in China, EU & the U.S.", accessed 31 May 2025, https://pernot-leplay.com/ai-regulation-china-eu-us-comparison/; Sheehan, Matt, "China's AI Regulations and How They Get Made", Carnegie Endowment for International Peace, July 2023, https://carnegie-production-assets.s3.amazonaws.com/static/files/202307-Sheehan_Chinese%20AI%20gov-1.pdf; Sheehan, Matt, "What the U.S. Can Learn from China about Regulating AI", Foreign Policy, 12 September 2023, https://foreignpolicy.com/2023/09/12/ai-artificial-intelligence-regulation-law-china-us-schumer-congress/.

that is responsive to the rapidly evolving nature of AI technologies. One advantage of this approach is that legislation for such pressing issues can be rolled out relatively swiftly without waiting for an overarching legislation to be put in place.[284]

3.2 The pro-innovation approach of the UK and US at federal level

The EU's approach can be contrasted with the approaches taken in the UK and the US at the federal level, which might be characterised as "pro-innovation". Both the UK Government and the US Federal Government have at present opted against introducing comprehensive AI legislation and at the same time are seeking to accelerate AI innovation and adoption across the public and private sectors.

The UK Government has opted against introducing comprehensive AI legislation. Prof Oliver Quick, Professor of Health Law and Policy at the University of Bristol in the UK, traces the origins of the UK's approach to two key documents: a report by the House of Lords Select Committee on AI[285] and the Government National AI Strategy.[286] The "pro-innovation" stance has subsequently been reinforced by consecutive Governments, notably via the introduction of a set of cross-sector principles for AI regulation[287] and an AI Opportunities Action Plan.[288]

Prof Quick described the UK's approach to AI regulation as "light touch" and "based on soft compliance". The five cross-sector principles – of safety, security and robustness; appropriate transparency and explainability; fairness; accountability and governance; and contestability and redress – are based on those previously developed by the OECD.[289] However, in Prof Quick's assessment, these are less exacting by reason of the omission of references to privacy, human rights and societal wellbeing. For Prof Quick, it is unclear how this principles-based approach to regulation will be delivered and how these principles will be developed into enforceable requirements. The benefits associated with this approach, namely, promoting innovation and developing a flexible and adaptive system that can be made to fit specific sectors and fast-developing technologies, therefore, need to be balanced alongside persisting concerns around unaddressed risks and a lack of enforcement.

A clear argument can be made for the regulation of AI being retained within specific sectors, with oversight delegated to existing regulators. Analysis of the EU AI Act indicates the complex consequences of overlapping regulatory systems. One potential issue with this approach, however, is that some of the matters that ought to be regulated might not fall within the purview of any of the existing regulators. In the UK, a number of separate but inter-related bodies are responsible for regulating different aspects of healthcare, including the Medicines and Healthcare products Regulatory Agency (MHRA), the National Institute for Health and Care Excellence (NICE), the Health Research Authority (HRA), and the Care Quality Commission (CQC). The MHRA plays a particularly important role in regulating AI in the context of medical devices under medical device

---

[284] *Ibid.*
[285] UK House of Lords Select Committee on Artificial Intelligence, "AI in the UK: Ready, Willing and Able? Report of Session 2017–19", HL Paper 100, 2018, https://committees.parliament.uk/committee/376/artificial-intelligence-committee/news/94648/uk-can-lead-the-way-on-ethical-ai-says-lords-committee/.
[286] UK Government, "National AI Strategy", updated 18 December 2022, https://www.gov.uk/government/publications/national-ai-strategy.
[287] Department for Science, Innovation and Technology of the UK Government, "Implementing the UK's AI Regulatory Principles: Initial Guidance for Regulators", February 2024, https://assets.publishing.service.gov.uk/media/65c0b6bd63a23d0013c821a0/implementing_the_uk_ai_regulatory_principles_guidance_for_regulators.pdf.
[288] Secretary of State for Science, Innovation and Technology of the UK Government, "Independent Report: AI Opportunities Action Plan", 13 January 2025, https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan.
[289] Organisation for Economic Co-operation and Development (OECD), "AI Principles", accessed 31 May 2025, https://www.oecd.org/en/topics/sub-issues/ai-principles.html

law, notably the UK Medical Devices Regulations 2002 [290]. MHRA, the Food and Drug Administration (FDA) of the US and Health Canada jointly issued "Good Machine Learning Practice for Medical Device Development: Guiding Principles"[291], "Predetermined Change Control Plans for Machine Learning-enabled Medical Devices: Guiding Principles"[292] and "Transparency for Machine Learning-enabled Medical Devices: Guiding Principles" [293]. MHRA has produced extensive guidance on the regulation of software as a medical device (e.g. guidance titled "Software and Artificial Intelligence (AI) as a Medical Device" [294]) and is perhaps best-placed to assess some of the risks posed by applications of AI in healthcare. In a policy paper titled "Impact of AI on the Regulation of Medical Products: Implementing the AI White Paper Principles" published in 2024, MHRA described its approach to regulating AI in the context of medical devices and a regulatory reform under way.[295] A regulatory sandbox called "AI Airlock" was launched by the MHRA.[296] As pointed out by the MHRA:

> AI Airlock is a regulatory "sandbox", a type of study where manufacturers can explore how best to collect evidence that could later be used to support the approval of their product. This is done under MHRA supervision in a virtual or simulated setting. Doing so will help the manufacturer and the MHRA better understand the challenges of regulating AI in medical devices, leading to a more bespoke and enabling regulatory framework …[297]

Nevertheless, the UK Government's emphasis on pro-innovation might present challenges for sectoral regulators such as the MHRA. Prof Quick suggests there is an inevitable trade-off between innovation and safety, and the current direction of travel in the UK "appears to put safety second".

The US Federal Government, too, has adopted a pro-innovation approach.

At the state level (in some of the states, e.g. Colorado) and local level in the US, laws have been passed to regulate AI in various contexts. [298] However, on 22 May 2025, the US House of

[290] The Medical Devices Regulations 2002 (SI 2002 No. 618, as amended) of the UK, https://www.legislation.gov.uk/uksi/2002/618/contents.

[291] Medicines and Healthcare products Regulatory Agency of the UK, US Food and Drug Administration, and Health Canada, "Good Machine Learning Practice for Medical Device Development: Guiding Principles", 27 October 2021, https://www.gov.uk/government/publications/good-machine-learning-practice-for-medical-device-development-guiding-principles/good-machine-learning-practice-for-medical-device-development-guiding-principles.

[292] Medicines and Healthcare products Regulatory Agency of the UK, US Food and Drug Administration, and Health Canada, "Predetermined Change Control Plans for Machine Learning-enabled Medical Devices: Guiding Principles", 24 October 2023, https://www.gov.uk/government/publications/predetermined-change-control-plans-for-machine-learning-enabled-medical-devices-guiding-principles.

[293] Medicines and Healthcare products Regulatory Agency of the UK, US Food and Drug Administration, and Health Canada, "Transparency for Machine Learning-enabled Medical Devices: Guiding Principles", 13 June 2024, https://www.gov.uk/government/publications/machine-learning-medical-devices-transparency-principles.

[294] Medicines and Healthcare products Regulatory Agency of the UK, "Guidance: Software and Artificial Intelligence (AI) as a Medical Device", updated 3 February 2025, https://www.gov.uk/government/publications/software-and-artificial-intelligence-ai-as-a-medical-device/software-and-artificial-intelligence-ai-as-a-medical-device.

[295] Medicines and Healthcare products Regulatory Agency of the UK, "Impact of AI on the Regulation of Medical Products: Implementing the AI White Paper Principles", 30 April 2024, https://www.gov.uk/government/publications/impact-of-ai-on-the-regulation-of-medical-products.

[296] Medicines and Healthcare products Regulatory Agency of the UK, "AI Airlock: the Regulatory Sandbox for AIaMD", updated 30 May 2025, https://www.gov.uk/government/collections/ai-airlock-the-regulatory-sandbox-for-aiamd.

[297] Medicines and Healthcare products Regulatory Agency of the UK, "Press Release: MHRA Trials Five Innovative AI Technologies as Part of Pilot Scheme to Change Regulatory Approach", 4 December 2024, https://www.gov.uk/government/news/mhra-trials-five-innovative-ai-technologies-as-part-of-pilot-scheme-to-change-regulatory-approach.

[298] White & Case, "Automated Decision Making Emerges as an Early Target of State AI Regulation", 7 March 2025, https://www.whitecase.com/insight-alert/automated-decision-making-emerges-early-target-state-ai-regulation; White & Case, "From California to Kentucky: Tracking the Rise of State AI Laws in 2025", 27 May 2025, https://www.whitecase.com/insight-alert/california-kentucky-tracking-rise-state-ai-laws-2025.

Representatives passed a set of provisions in a budget reconciliation package, inclusive of what is described as "a ban on state and local enforcement of AI legislation and regulations" and a "[t]en-year moratorium on AI regulation".[299] At the time of writing, it remains to be seen whether the bill will be passed by the US Congress.

The US Federal Government, under the current administration, has taken steps to reverse protective regulatory measures at the federal level.[300] Although comprehensive AI legislation has not been introduced via Congress, presidential executive orders have sought to steer the regulation of AI in the US. Via Executive Order No. 14110[301], for example, former President Joe Biden directed agencies, including the US Department of Health and Human Services, to take action to promote safety and security, alongside innovation, in the development and use of AI. This was rescinded by President Donald Trump at the start of his current term, signalling a policy shift that seems likely to result in deregulation in support of AI industry growth.[302]

Prof Boris Babic, Associate Professor in the Department of Philosophy and (by courtesy) Faculty of Law at the University of Hong Kong, and Prof Sara Gerke, Associate Professor of Law at the University of Illinois in the US, discussed the US medical device law. This, as Prof Boris Babic explained, has its modern origins in the 1976 Medical Device Amendments to the 1938 Federal Food, Drug, and Cosmetic Act (FDCA). US FDA regulates all AI products that meet the definition of a medical device as outlined in section 201(h)(1) of the FDCA. As with other medical devices, AI medical devices must be subject to a pre-market submission process and, for those that enter the market, post-market surveillance. Prof Gerke reported that the majority of AI-based medical devices whose marketing had been permitted by the FDA went through a pathway called the 510(k) pathway (Premarket Notification). As pointed out by Prof Gerke, in the 510(k) pathway, applicants need to show that their medical devices are "substantially equivalent" to a legally marketed device and the 510(k) pathway may or may not require any clinical evidence to be provided.

Both Prof Babic and Prof Gerke expressed concerns with the existing system for regulating AI medical devices in the US. Prof Gerke proposed that AI medical devices require specific labelling requirements that go beyond those required for standard medical devices and explained that inadequate labelling of AI medical devices poses a risk of harm to patients as a result of biased care or pointless treatment. Prof Babic's research also reveals challenges for the post-market surveillance of AI medical devices: close inspection of the FDA's database of adverse events associated with medical devices reveals significant gaps and misclassification in the data. The extent of missing data is significantly higher among AI medical devices compared with other medical devices. As a result of the gaps and misclassification, the data offered limited insight into the risks posed by these AI medical devices.

---

[299] DLA Piper, "Ten-year Moratorium on AI Regulation Proposed in US Congress", 22 May 2025, https://www.dlapiper.com/en/insights/publications/ai-outlook/2025/ten-year-moratorium-on-ai; Hogan Lovells, "U.S. House of Representatives Passes Proposal to Prohibit Enforcement of State AI Laws for 10 years", 22 May 2025, https://www.hoganlovells.com/en/publications/us-house-of-representatives-passes-proposal-to-prohibit-enforcement-of-state-ai-laws-for-10-years.

[300] The White House, "Fact Sheet: President Donald J. Trump Takes Action to Enhance America's AI Leadership", 23 January 2025, https://www.whitehouse.gov/fact-sheets/2025/01/fact-sheet-president-donald-j-trump-takes-action-to-enhance-americas-ai-leadership/.

[301] Federal Register, "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence: A Presidential Document by the Executive Office of the President on 11/01/2023", https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.

[302] Kirkland & Ellis, "Considering the Future of AI Regulation on the Health Sector", 5 March 2025, https://www.kirkland.com/publications/article/2025/03/considering-the-future-of-ai-regulation-on-health-sector.

4. Status of regulation in other jurisdictions

The EU has put in place a comprehensive and cross-sector AI law, the EU AI Act, prioritising safeguards to promote safety, privacy, and fairness. In contrast, the UK and, at the federal level, the US have adopted a more pro-innovation stance. However, these two extremes do not capture the full spectrum of approaches among the jurisdictions considered at the conference. Some of these jurisdictions occupy a middle ground, relying relatively heavily on existing legislation and bolstering it with a combination of AI-specific regulation and guidance.

Macau & Mainland China

Whilst Mainland China has not introduced an all-encompassing piece of AI legislation akin to the EU AI Act, Mainland China has adopted a proactive approach. The Government is actively promoting the growth of AI, including in healthcare, and at the same time has enacted AI-specific regulations, such as regulations to address concerns around recommendation algorithms, deep synthesis, and generative AI technologies (e.g. labelling of AI-generated content).[303] The current AI regulatory framework of Mainland China is described as a framework that "combines high-level national plans, targeted regulations for specific AI applications … extensive technical standards … and application of existing laws".[304]

Dr Ji Ping, Vice Director of the Clinical Research Institute at Shenzhen Peking University-Hong Kong University of Science and Technology Medical Center, discussed the registration/ filing and classification of AI medical devices, which is overseen by the National Medical Products Administration (NMPA) of the People's Republic of China. NMPA has issued specific guidance on AI medical devices, such as "Guiding Principles for the Registration Review of AI Medical Devices"[305]. Dr Ji noted the landscape for AI development and discussed the current status and barriers. Dr Ji highlighted that there were difficulties around obtaining and sharing data and the possibility of over-reliance on or distrust of AI technology. The difficulties around obtaining and sharing data may be particularly inhibiting given the data-intensive nature of AI development. This challenge was also emphasised by Mr Zhangyu Wang, PhD student, and Prof Li Du, Associate Professor, both in the Faculty of Law at the University of Macau, who discussed the challenges in respect of the use and cross-border transfer of data given the regulatory landscape in Macau and Mainland China.

Middle East

Dr Barry Solaiman, Assistant Professor of Law at Hamad Bin Khalifa University in Qatar, discussed the regulatory landscape of Qatar, Saudi Arabia, and the United Arab Emirates (UAE). Dr Solaiman highlighted the medical device guidance published by the Saudi Food and Drug Authority in Saudi Arabia ("MDS-G010: Guidance on Artificial Intelligence and Machine Learning Technologies Based Medical Devices"[306]) and two UAE documents ("Policy on Use of Artificial Intelligence (AI) in the

[303] Herbert Smith Freehills Kramer, "AI Tracker: Mainland China", updated 23 May 2025, https://www.herbertsmithfreehills.com/insights/reports/searchlight-ai/prc.
[304] *Ibid.*
[305] Center for Medical Device Evaluation of the National Medical Products Administration of the Government of the People's Republic of China (中华人民共和国国家药品监督管理局医疗器械技术审评中心), "Guiding Principles for the Registration Review of AI Medical Devices" (人工智能医疗器械注册审查指导原则), 2022.
[306] Saudi Food and Drug Authority, "Guidance on Artificial Intelligence (AI) and Machine Learning (ML) Technologies Based Medical Devices (MDS-G010) Version 1.0", 29 November 2022, https://www.sfda.gov.sa/sites/default/files/2023-01/MDS-G010ML.pdf.

Healthcare Sector of the Emirate of Abu Dhabi" issued by the Department of Health of Abu Dhabi and a policy on AI in the health sector issued by Dubai Health Authority). He also shared about a set of guidelines titled "Research Guidelines for Healthcare AI Development", which was the output of a Qatar research grant.[307]

Canada

In Canada, the approach to AI regulation at the federal level is currently uncertain. In June 2022, the Government of Canada tabled the Artificial Intelligence and Data Act (AIDA) as part of Bill C-27, the Digital Charter Implementation Act, 2022.[308] The purpose of AIDA was to introduce specific measures for regulating AI in Canada, but its progress through Parliament stalled and it could not be passed prior to the 2025 Canadian election. It was predicted that the Conservative Party, if it won the election, would "want to narrow the application of any future AI legislation and/or clearly define, at the outset, how the law would be implemented and enforced". [309] However, the Liberal Party was re-elected and it remains to be seen whether the Liberal Party will revive AIDA.[310]

At the conference, Yuan Stevens, Academic Associate in the Centre of Genomics and Policy at McGill University in Canada, described the regulatory landscape in Canada, focusing on the governance of medical devices under the Medical Devices Regulations of Canada ("CA MDR") and related guidance documents. In 2019, Health Canada issued "Guidance Document: Software as a Medical Device: Definition and Classification", which clarifies which products qualify as Software as a Medical Device, as well as how Software as a Medical Device is classified under the CA MDR. [311] In 2025, Health Canada issued "Pre-market Guidance for Machine Learning-enabled Medical Devices", which sets out supporting information to consider for demonstration of the safety and effectiveness of machine learning-enabled medical devices under CA MDR by manufacturers (a) for the purposes of applications for or amending a class II, III or IV medical device licence or (b) at any point in the device lifecycle (class I to class IV).[312]

[307] Solaiman, Barry, Ghaly, Mohammed, Househ, Mowafa, *et al.*, "Research Guidelines for Healthcare AI Development Version 1.0", Hamad Bin Khalifa University, April 2025, developed under the research grant "Artificial Intelligence for Precision Medicine & Health Technologies: Developing a Regulatory Framework for Qatar and the Middle East" (HBKU-SRO-TGA-VPR-TG01-001), http://dx.doi.org/10.13140/RG.2.2.10590.14402.
[308] Government of Canada, "The Artificial Intelligence and Data Act (AIDA) – Companion Document", modified 31 January 2025, https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document.
[309] Baker McKenzie, "An Election Is Looming – The Future of Canadian AI Legislation", 17 April 2025, https://canada-insights.bakermckenzie.com/2025/04/17/an-election-is-looming-the-future-of-canadian-ai-legislation/.
[310] Dentons, "Artificial Intelligence Trends to Watch in 2025: Regulation of AI", 4 February 2025 https://www.dentons.com/en/insights/newsletters/2025/january/23/global-regulatory-trends-to-watch/dentons-canadian-regulatory-trends-to-watch-in-2025/artificial-intelligence-trends-to-watch-in-2025.
[311] Health Canada, "Guidance Document: Software as a Medical Device (SaMD): Definition and Classification", 18 December 2019, https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance-document.html.
[312] Health Canada, "Pre-market Guidance for Machine Learning-enabled Medical Devices", 5 February 2025, https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices.html.

5. The real-world application of medical AI: lessons from assisted reproduction

While most speakers discussed the regulation of a broad category of medical AI, the appropriate nature and content of regulatory requirements will need to be sensitive to the context of application and specific use case. Speakers from Hong Kong and Macau provided an excellent example of this by exploring regulatory challenges for AI through the lens of assisted reproductive technologies (ART). AI is now used in embryo selection for the purposes of in vitro fertilization (IVF).

Prof Calvin Wai-Loon Ho, Associate Professor of Law at Monash University in Australia, argued that it is possible that the current regulatory regime in Hong Kong does not adequately support the evaluation of innovation in AI-enabled assisted reproduction, and flagged the need to establish a regulatory environment that facilitates evidence generation for AI-based medical interventions.

Prof Ho explained as follows. In IVF, decision-making is required at different stages of embryonic development prior to implantation to ensure a successful pregnancy, and many of these decisions are very subjective and will vary greatly depending on clinical experience. Now, data-driven approaches and AI technology are being incorporated in order to facilitate optimal, consistent and objective decision-making, and to drive individualised treatment. These approaches range from "human-in-the-loop" AI clinical decision support for embryo selection to algorithmic drug dosing tools.[313] As pointed out by Prof Ho, there is limited evidence to show that the use of AI in embryo selection greatly advances clinical outcome and a study suggests that an AI-enabled approach to embryo selection might bring about almost no improvement in outcome in comparison with the conventional approach.[314] Prof Ho expressed concern that, despite the above, given the "hype" around these new AI-enabled approaches, patients are prepared to pay more for them. He noted that Hong Kong does not currently have a regulatory framework that supports the production of reliable evidence to guide policy and patient decision-making, apart from generic clinical trial regulatory guidelines.

These challenges are further compounded by regulatory gaps in data governance, which is fundamental to the development and validation of AI in healthcare. Mr. Zhangyu Wang and Prof Li Du (University of Macau) raised the complex interplay of regulatory instruments relevant to the use of and transfer of data for the purposes of research to develop AI-enabled ART. They argued that, in Mainland China, the relevant legal rules — particularly those concerning human genetic data, medical record localisation, and restrictions on cross-border data transfer — might constitute hurdles to be overcome for collaborative research. These restrictions might make cross-jurisdictional AI development in the ART space more difficult, limiting the potential for pooled datasets and collaborative validation efforts that are vital for improving performance and generalisability in medical AI.

6. Embedding ethical considerations into AI governance

Conference speakers emphasised the importance of embedding ethical principles into governance frameworks for medical AI. This is important, not only to ensure that AI is deployed safely, but to demonstrate trustworthiness to clinicians using AI tools and to patients who might need to accept AI as part of clinical processes. Although not fundamentally exceptional, the autonomous nature of AI, its lack of interpretability and susceptibility to embedding biases raise additional concerns beyond those generated by other emerging medical technologies. Speakers explored a number of these

[313] Hanassab, Simon, Abbara, Ali, Yeung, Arthur C. *et al*., "The Prospect of Artificial Intelligence to Personalize Assisted Reproductive Technology." *npj Digital Medicine* 7 (2024): 55. https://doi.org/10.1038/s41746-024-01006-x.
[314] Illingworth, Peter J., Venetis, Christos, Gardner, David K. *et al*., "Deep Learning Versus Manual Morphology-based Embryo Selection in IVF: a Randomized, Double-blind Noninferiority Trial", *Nature Medicine* 30, no. 11 (2024): 3114, http://doi.org/10.1038/s41591-024-03166-5.

concerns, highlighting potential harms and identifying key questions and considerations for policymakers.

Amongst these concerns was the potential for AI to be biased and amplify and systematise biases, possibly exacerbating health inequalities (as discussed by Mr Donald Mak, Prof Rachel Sterken, Ms Tanya Brigden, and Prof Brian Wong). This has been exemplified by several AI systems that have shown the ability of algorithms to systematically misrepresent and exacerbate health problems in underrepresented groups. Governance mechanisms to mitigate against biases are crucial. This could include requirements for transparent and clear reporting of limitations and biases of datasets, and incentivising the curation and use of datasets for AI systems that are diverse, inclusive, and promote AI generalisability.

Another concern was raised around the potential for AI to demonstrate persuasive and manipulative capabilities, leading to the erosion of trust. Prof Rachel Sterken, Associate Professor in the Department of Philosophy at the University of Hong Kong, argued that AI systems, in the way outputs are given and perceived, may unduly influence medical professionals into rejecting their own diagnosis and accepting that of the AI. Prof Sterken emphasised the need to be wary of a "computer-knows-best" mentality, where machines dictate treatment options based on their programmed priorities rather than the patient's individual preferences, a concern shared by Prof Brian Wong, Assistant Professor in the Department of Philosophy at the University of Hong Kong, who cautioned that "what the AI construes to be 'best'", may not in fact be what the patient views as "best", let alone what is ultimately "best".

These potential harms highlight the need for governance mechanisms, such as human oversight and explainability requirements. Transparency and explainability were identified as central ethical requirements. Many AI products currently operate as "black boxes", making decisions that are not easily interpretable by human users. This lack of explainability undermines medical professionals' ability to rely on AI systems and complicates the detection of errors or biases. There was extensive discussion among participants around explainability, whether to exclude black-box AI from healthcare provision and the situations in which explainable decisions are essential. Although explainability is undeniably an important safeguard, the extent to which the output of AI needs to be explainable may be context-dependent, taking into consideration how it will be used in the clinical pathway, the degree of risk to the patient and evidence of clinical utility.

Human oversight is a potentially important safeguard against harms arising as a result of bias or black box decision-making. However, AI has now been demonstrated to outperform humans in some tasks conventionally requiring clinical judgement, raising questions around the degree of human oversight that is reasonable and proportionate where patients may benefit from greater automation. Dr Pete Mills, Director at the PHG Foundation, put forward a number of different ways that the interactions between AI and humans could be conceptualised in healthcare. In doing so, he raised some of the challenges with "human-AI teaming", whilst still acknowledging that reliable autonomous agents have considerable potential to contribute to medical practice. He emphasised the concept of trust as an important part of clinical practice, noting that trust has an inherent moral component which cannot apply to machines that lack moral agency.

Cognisant of these possible harms, many jurisdictions already have ethical governance codes, guidelines and frameworks in place.[315] [316] [317] [318] [319] [320] Despite their differences, they collectively demonstrate a shared recognition of the ethical risks posed by AI and a commitment to embedding human-centred values into AI governance. The challenge lies in translating these high-level principles into enforceable standards and operational guidance, particularly in complex fields like medical AI.

Prof Wong suggested that, in generating regulations and protocol governing AI usage, as well as conceptualising broader policies on AI adoption in medical contexts, policymakers should devise a set of discretionary principles and that discretionary duties and the rights to which they correlate shall apply to those who use AI, train AI, or advocate or lobby for greater incorporation of AI.

7. Conclusions

The range of developments described in this report demonstrates the growing demand for regulatory responses to medical AI. Although a diversity of current approaches was described by the speakers, there was broad agreement around a number of considerations and values that should underpin responsible AI governance. The following key themes and learnings from the conference may assist policymakers in Hong Kong and other jurisdictions as they navigate this complex and evolving field.

An adaptive approach to AI regulation while safeguarding against novel potential harms

An adaptive and flexible approach to medical AI regulation offers significant advantages in responding to the rapid pace of technological advancements. As argued by Prof Timo Minssen, the last-minute inclusion of provisions addressing generative AI in the EU AI Act underscored the difficulties of creating static regulations in a dynamic environment.

Regulatory frameworks that incorporate flexibility and are designed to evolve with emerging technologies are better suited to keep pace with these changes. However, this flexibility must be anchored by clear mechanisms for implementation and enforcement. This may necessitate the adoption of a risk-based approach that allows for ongoing adjustments and improvements based on real-world experiences and outcomes. As mentioned by Prof Minssen, regulatory sandboxes, as provided for in Article 57 of the EU AI Act, can be an important element in this endeavour.

Despite their differences, both the EU's and UK's approaches seek to promote safe and responsible AI technologies, whilst also creating a regulatory environment conducive to innovation. However,

---

[315] Digital Policy Office of the Government of the HKSAR, "Ethical Artificial Intelligence Framework (Customised Version for General Reference by Public) Version: 1.4", updated July 2024, https://www.digitalpolicy.gov.hk/en/our_work/data_governance/policies_standards/ethical_ai_framework/.

[316] National New Generation Artificial Intelligence Governance Professional Committee (国家新一代人工智能治理专业委员会) of the People's Republic of China, New Generation Artificial Intelligence Governance Principles (新一代人工智能治理原则), 2019.

[317] National New Generation Artificial Intelligence Governance Professional Committee (国家新一代人工智能治理专业委员会) of the People's Republic of China, New Generation Artificial Intelligence Code of Ethics (新一代人工智能伦理规范), 2021.

[318] Ministry of Foreign Affairs of the People's Republic of China (中华人民共和国外交部), Position Paper of China on Strengthening Ethical Governance of Artificial Intelligence (中国关于加强人工智能伦理治理的立场文件), 2022.

[319] Government of Canada, "Pan-Canadian AI for Health (AI4H) Guiding Principles", modified 30 January 2025, https://www.canada.ca/en/health-canada/corporate/transparency/health-agreements/pan-canadian-ai-guiding-principles.html.

[320] OECD.AI, "OECD AI Principles Overview", updated 2024, https://oecd.ai/en/ai-principles.

achieving this dual objective poses significant challenges and requires careful scrutiny to ensure the scales are not tipped too far in either direction.

Aligning AI governance with Medical Device Regulation

The Department of Health in Hong Kong has recently established the Preparatory Office for the Hong Kong Centre for Medical Products Regulation (CMPR) in 2024. The specific work of the Preparatory Office includes, among others, studying and planning a regulatory and approval regime for drugs and medical devices[321]. This seems to initiate a shift away from the voluntary MDACS and towards developing statutory powers over the regulation of medical devices, including qualifying medical AI products. In doing so, Hong Kong has the opportunity to design a regulatory system that learns from the varied approaches being taken to the regulation of AI medical devices around the world. As in other jurisdictions with existing medical device regulations, it will be important to consider how these align with wider AI governance initiatives, in order to identify possible duplication of requirements or gaps.

Governance across the life cycle

It is important to govern AI across the total lifecycle, from research and development, during training and validation, and then throughout deployment. The capacity that AI demonstrates to adapt and evolve post-deployment is one of the key challenges for governing medical AI. Assessing the safety and efficacy of medical AI in clinical settings prior to deployment can help mitigate risks. Conventional approaches to post-market surveillance may also need to be augmented to monitor specific safety concerns for AI medical devices, including concept drift, covariate shift and algorithmic stability. As suggested by Prof Babic, one possible measure to address the issue of concept drift is to mandate manufacturers to report regularly, in relation to their AI medical devices, about any significant updates to training data and any substantial amendments to deployment conditions.[322] Prof Babic also made suggestions on how the reporting requirements imposed on manufacturers of AI medical devices could be modified in order to address the issues of covariate shift and algorithmic stability.[323] It may also be helpful to mandate the use of labels with an "eye-popping" design, similar to "nutrition facts" labels, to provide users, etc. with relevant information, such as information about the data sets (e.g. general ethnicity breakdown), validation and model performance (in particular, cross-site performances) in accordance with labelling standards that are tailored to AI medical devices, as advocated by Prof Gerke.[324]

Regulating products and systems

AI requires regulators to expand their view beyond products to systems. The safety risks posed by AI medical devices are highly dependent on user interaction and cannot be identified solely via product monitoring. Prof Gerke's analogy of hospitals "hiring" AI medical devices, rather than "buying" them, helps to convey the iterative and collaborative approach to device development and deployment that is required. Adopting a systems approach can facilitate more effective collaboration between regulators and other stakeholders, such as developers, manufacturers, and users. In Shenzhen, China, the requirement to establish a mechanism for sharing responsibilities between actors imposed by the

[321] Department of Health of the Government of the HKSAR, "Preparatory Office for the Hong Kong Centre for Medical Products Regulation", revised 5 June 2024, https://www.dh.gov.hk/english/main/main_pocmpr/main_pocmpr.html.
[322] Babic, Boris, I. Glenn Cohen, Ariel Dora Stern, Yiwen Li, and Mellisa Ouellet, "A General Framework for Governing Marketed AI/ML Medical Devices", *npj Digital Medicine* 8 (2025): 328. https://doi.org/10.1038/s41746-025-01717-9.
[323] *Ibid.*
[324] Gerke, Sara, "'Nutrition Facts Labels' for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards", *The George Washington Law Review* 91, 1 (2023): 79. https://ssrn.com/abstract=4404252.

"Notice on the Safety Management of 'AI+Healthcare' Application", referred to by Dr Ji Ping, provides an example of this being implemented.

Global coordination

It will be vital to pay attention to global developments in AI regulation, particularly given cross-border efforts to develop the technology. Since the conference, developments in the US and Canada have demonstrated that AI regulation is sensitive to political change. Monitoring and engaging with AI governance at a global level will be important for understanding emerging trends and the implications for interaction between jurisdictions. The EU AI Act, for example, has extra-territorial scope. Article 2(1) of the EU Act provides that this Regulation applies to, among others, "providers placing on the market or putting into service AI systems or placing on the market general-purpose AI models in the Union, irrespective of whether those providers are established or located within the Union or in a third country" and "providers and deployers of AI systems that have their place of establishment or are located in a third country, where the output produced by the AI system is used in the Union". It is clear from Article 3(9) of the EU AI Act that "placing on the market" refers to placing on the Union market. If a provider or deployer established or located in Hong Kong falls within the scope of such extra-territorial provisions of the EU AI Act (e.g. output produced by the AI system is used in the EU), the provider/ deployer will be required to adhere to the applicable requirements of the EU AI Act. Data localisation and barriers to cross-border data transfers can also impede AI research, as explained by Mr Zhangyu Wang and Prof Li Du.

Regulation embedding ethical values

Finally, we are at an early stage in understanding the power and impact of AI technology in healthcare. Regulation and governance have an important role to play in order to safeguard against potential harms, to provide clarity and consistency for stakeholders, and ultimately to secure the trust and confidence of patients and the public. It is crucial to embed ethical values into the regulation of AI. As AI is integrated into the clinical process, adopters should, where necessary, ensure these tools are validated for use among local populations.

This will not be a one-off process, and policymakers will need to engage in an ongoing dialogue with AI developers, healthcare providers, patients, publics, academics, etc., to ensure that the regulatory framework keeps pace with technical developments, evolving societal expectations and developing evidence on the ways in which AI is impacting the practice of and values underpinning healthcare.

Appendix 1

Conference Programme

The Programme –
Day 1
Sessions 1 and 2 (by invitation only)

15 November 2024, Friday
9:00 – 13:15 (Hong Kong Time)
11/F, Cheng Yu Tung Tower, Centennial Campus, HKU

9:00 – 9:05	**Welcome and introduction** **Prof Gilberto Leung** Co-Director, Centre for Medical Ethics and Law Tsang Wing-Hing Professor in Clinical Neuroscience and Clinical Professor, LKS Faculty of Medicine The University of Hong Kong & President, Hong Kong Academy of Medicine Hong Kong, China
9:05 – 9:10	**Speech by Guest of Honour** **Dr LEE Ha Yun, Libby, JP** Under Secretary for Health, Health Bureau, Government of the Hong Kong Special Administrative Region (HKSAR), China
9:10 – 9:15	**Photo session**
Session 1: Governance of Healthcare AI in the EU and the UK **Chair:** **Prof Eric C. Ip** Co-Director of the Centre for Medical Ethics and Law & Professor of the Faculty of Law The University of Hong Kong, Hong Kong, China	
9:15 – 9:45	**Shaping tomorrow's healthcare: The European Union AI Act and its implications in medicine** **Prof Vera Lúcia Raposo** Associate Professor of Law and Technology, NOVA School of Law, Portugal
9:45 – 10:15	**The European Health Data Space: Realistic ambitions for health research and innovation?**

	Dr Colin Mitchell Head of Humanities, PHG Foundation, University of Cambridge, UK
10:15 – 10:35	**The UK approach to regulating AI: Innovation first, safety second?** **Prof Oliver Quick** Professor of Health Law and Policy, University of Bristol Law School, UK
10:35 – 10:45	**Intermission**
Session 2: Governance of Medical AI: the Local Situation **Chair:** Prof Gilberto Leung Co-Director, Centre for Medical Ethics and Law Tsang Wing-Hing Professor in Clinical Neuroscience and Clinical Professor, LKS Faculty of Medicine The University of Hong Kong & President, Hong Kong Academy of Medicine Hong Kong, China	
10:45 – 11:05	**Opportunities and challenges of applying AI for medical applications** **Mr Donald Mak** Deputy Commissioner for Digital Policy (Data Governance) Digital Policy Office, Innovation, Technology and Industry Bureau, Government of the HKSAR, China
11:05 – 11:35	**AI-enabled assisted reproduction: regulatory and ethical issues in Mainland China and Macau & Hong Kong perspective on AI-enabled assisted reproduction** **Prof Stephen Li Du** Associate Professor, Faculty of Law, University of Macau, Macau, China **Prof Calvin Wai-Loon Ho** Research Fellow, Centre for Medical Ethics and Law, The University of Hong Kong, Hong Kong, China Associate Professor, Monash Law School**,** Monash University, Australia

11:35 – 11:55	**Ethics governance on medical AI research and development in Shenzhen, China** **Dr JI Ping** Vice Director, Clinical Research Institute, Shenzhen Peking University-Hong Kong University of Science and Technology Medical Center, China
11:55 – 12:15	**AI for medical innovation – from medical device's perspective** **Mr LAM Kam Chun, Tommy** Senior Electronics Engineer (Medical Device) Medical Device Division, Department of Health Government of the HKSAR, China
12:15 – 13:15	**Roundtable Discussion** **Prof Vera Lúcia Raposo, Dr Colin Mitchell, Prof Oliver Quick, Mr Donald Mak, Prof Stephen Li Du, Prof Calvin Wai-Loon Ho, Dr JI Ping, Mr LAM Kam Chun, Tommy** Chairs: Prof Gilberto Leung and Prof Eric C. Ip
13:15 – 14:30	**Lunch (for invitees only)**

The Programme –
Day 1
Sessions 3 & 4 (by invitation only)

15 November 2024, Friday
14:30 – 17:45 (Hong Kong Time)
11/F, Cheng Yu Tung Tower, Centennial Campus, HKU

Session 3: More on AI Governance in the EU and Canada **Chair:** Dr Pete Mills Director, PHG Foundation, University of Cambridge, UK	
14:30 – 14:50	**Ensuring a trustworthy use of medical AI through regulation - the European perspective** **Prof dr Tom Goffin** Associate Professor of Health Law, Metamedica, Ghent University, Belgium
14:50 – 15:10	**Navigating the EU AI Act and the Medical Device Regulation (MDR): Implications for regulated digital medical products** **Prof Timo Minssen** Professor of Law Managing Director and Founder of Center for Advanced Studies and the Collaborative Research Programme in Biomedical Innovation Law (CeBIL) University of Copenhagen, Denmark
15:10 – 15:30	**From code to care: Regulating AI in Canadian healthcare** **Yuan Y. Stevens** Academic Associate, Centre of Genomics and Policy, McGill University, Canada
15:30 – 16:00	**Q&A** **Prof dr Tom Goffin, Prof Timo Minssen, Yuan Y. Stevens** Chair: Dr Pete Mills
16:00 – 16:15	**Intermission**

Session 4: Regulating AI in the US and Beyond **Chair:** Dr Colin Mitchell Head of Humanities, PHG Foundation, University of Cambridge, UK	
16:15 – 16:35	**Regulating AI/ML in U.S. healthcare: challenges, Opportunities, and FDA's Evolving Framework** **Prof Sara Gerke** Associate Professor of Law and Richard W. & Marie L. Corman Scholar, College of Law, University of Illinois Urbana-Champaign, US
16:35 – 16:55	**Managing failure: How should we track and govern problems arising from medical AI/ML devices?** **Prof Boris Babic** Associate Professor, HKU Musketeers Foundation Institute of Data Science and Department of Philosophy Associate Professor (by courtesy), Faculty of Law The University of Hong Kong, Hong Kong, China
16:55 – 17:15	**A GCC approach towards regulating the lifecycle of healthcare AI** **Dr Barry Solaiman** Assistant Professor of Law, College of Law, Hamad Bin Khalifa University, Qatar, Doha
17:15 – 17:45	**Q&A** **Prof Sara Gerke, Prof Boris Babic and Dr Barry Solaiman** Chair: Dr Colin Mitchell
After session 4 ends	**Dinner (for invitees only)**

The Programme –
Day 2
Session 5

16 November 2024, Saturday
9:00 – 12:35 (Hong Kong Time)
11/F, Cheng Yu Tung Tower, Centennial Campus, HKU or Zoom

Session 5: Legal and Ethical Considerations of AI Governance in Healthcare	
9:00 – 9:10	**Introduction** **Prof Eric C. Ip** Co-Director, Centre for Medical Ethics and Law Professor, Faculty of Law The University of Hong Kong, Hong Kong, China
Chair before intermission: Dr Pete Mills Director, PHG Foundation, University of Cambridge, UK	
9:10 – 9:40	**Summary of discussion on Day 1** **Prof Gilberto Leung** Co-Director, Centre for Medical Ethics and Law Tsang Wing-Hing Professor in Clinical Neuroscience and Clinical Professor, LKS Faculty of Medicine The University of Hong Kong & President, Hong Kong Academy of Medicine Hong Kong, China **Prof Calvin Wai-Loon Ho** Research Fellow, Centre for Medical Ethics and Law, The University of Hong Kong, Hong Kong, China Associate Professor, Monash Law School**,** Monash University, Australia
9:40 – 10:00	**Role of AI in medical and public health policymaking** **Prof Brian Wong** Assistant Professor, Department of Philosophy, The University of Hong Kong, Hong Kong, China
10:00 – 10:20	**Medical AI assistants vs. Medical AI dictators**

	Prof Herman Cappelen Chair Professor, Department of Philosophy Director of AI & Humanity Lab and Co-director of ConceptLab The University of Hong Kong, Hong Kong, China
10:20 – 10:40	**Ethical considerations for the trustworthy use of medical AI** **Tanya Brigden** Senior Policy Analyst, PHG Foundation, University of Cambridge, UK
10:40 – 10:55	**Intermission**
Chair after intermission: Dr Colin Mitchell Head of Humanities, PHG Foundation, University of Cambridge, UK	
10:55 – 11:15	**The AI as expert: some epistemological and moral considerations** **Dr Pete Mills** Director, PHG Foundation, University of Cambridge, UK
11:15 – 11:35	**Emerging ethical challenges in medical AI: persuasion, manipulation and consent** **Prof Rachel Sterken** Associate Professor, Department of Philosophy Co-director of ConceptLab The University of Hong Kong, Hong Kong, China
11:35 – 12:35	**Roundtable Discussion** **Prof Gilberto Leung, Prof Calvin Wai-Loon Ho, Prof Herman Cappelen, Tanya Brigden, Dr Pete Mills, Prof Rachel Sterken** Chair: Dr Colin Mitchell

Appendix 2

Biographies

Prof Gilberto Leung
Co-Director, Centre for Medical Ethics and Law
Tsang Wing-Hing Professor in Clinical Neuroscience and Clinical Professor, LKS Faculty of Medicine
Director of the School of Clinical Medicine
The University of Hong Kong, Hong Kong Special Administrative Region, China

Biography:

Professor Gilberto Leung is a neurosurgeon, Clinical Professor, and Tsang Wing-Hing Professor in Clinical Neuroscience at the University of Hong Kong (HKU). He graduated from the University of London with M.B.B.S. in 1992 and joined HKU in 2005 where he obtained his MS, PhD, and MD. He currently serves as Director of the School of Clinical Medicine at the LKS Faculty of Medicine, Immediate-Past President of the Hong Kong Academy of Medicine, and Covenor of the Institute for Medical Advancement and Clinical Excellence in Hong Kong. He holds an LLB from the University of London and an LLM in Medical Law and Ethics with Distinction from the University of Edinburgh. He is Co-Director of the Centre for Medical Ethics and Law at HKU, and Co-Chairperson of the Professionalism and Ethics Committee at the Hong Kong Academy of Medicine.

Prof Eric C. Ip
Co-Director, Centre for Medical Ethics and Law
Professor, Faculty of Law
The University of Hong Kong, Hong Kong Special Administrative Region, China

Biography:

Prof Eric C. Ip is a public health bioethicist and Co-Director of the Centre for Medical Ethics and Law at the University of Hong Kong (HKU), where he is Professor of Law. He holds a Doctor of Philosophy in Socio-Legal Studies from the Faculty of Law, University of Oxford, and a Master of Bioethics from the Faculty of Medicine and Health, The University of Sydney, among other academic degrees. He has been awarded the Outstanding Young Researcher Award, University Research Output Prize, and Faculty Outstanding Teaching Award by HKU, and the Research Excellence Award and Young Researcher Award by The Chinese University of Hong Kong, where he served as Assistant Dean for Undergraduate Studies and Director of the LLB Programme.

Mr Donald Mak
Deputy Commissioner (Data Governance)
Digital Policy Office, Innovation, Technology and Industry Bureau
The Government of the Hong Kong Special Administrative Region

Biography:

As the Deputy Commissioner (Data Governance), Mr Donald Mak has been actively involved in promoting Government's agenda for innovation, technology adoption and public services improvement. He also played a key role in the implementation of a number of digital government infrastructures (including the government cloud infrastructure, big data analytics platform, shared blockchain platform, consented data exchange gateway) and the formulation and implementation of open data policy.

Mr Mak joined the Government in 1989. Since then he has assumed various important duties including formulation and implementation of facilitation strategy and initiatives for ICT industry development, data centre development, digital inclusion, ICT co-operation with the Mainland and overseas economies as well as the implementation and management of various government ICT systems and services.

Mr LAM Kam Chun, Tommy
Senior Electronics Engineer (Medical Device)
Medical Device Division, Department of Health
The Government of the Hong Kong Special Administrative Region

Biography:

Mr Tommy LAM is the Senior Engineer of the Medical Device Division in the Department of Health of the HKSAR Government. He is a corporate member in the Biomedical discipline of the Hong Kong Institution of Engineers and has over 16 years of sound experience in electronics and biomedical engineering, Mr Tommy LAM has previously worked in hospital engineering and innovation and technology at the Electrical and Mechanical Services Department, with a particular focus on artificial intelligence and digitalization projects. He is currently responsible for the implementation of the Medical Device Administrative Control System, with a view to facilitating transition to a long-term statutory regulatory framework for medical devices in Hong Kong.

Prof Calvin Wai-Loon Ho
Research Fellow, Centre for Medical Ethics and Law, The University of Hong Kong, Hong Kong, China
Associate Professor, Monash Law School, Monash University, Australia

Biography:

Prof Calvin Wai-Loon Ho is Associate Professor with the Monash Law School, a Research Fellow of the Centre for Medical Ethics and Law (CMEL) at the University of Hong Kong (HKU), a Fellow of PHG Foundation at the University of Cambridge, and a Fellow of the Royal Society for Public Health UK. He co-chairs the Expert Group of the World Health Organization (WHO) on the Ethics of Social Listening and Infodemic Management, and is an Ethics Board member of Médecins Sans Frontières (Doctors Without Borders).

Prior to his current appointment, Calvin was Associate Professor with HKU Faculty of Law, and Co-Director of CMEL, which is jointly run by the Faculty of Law and Faculty of Medicine at HKU. Before this, he was Assistant Professor with the Centre for Biomedical Ethics of the Yong Loo Lin School of Medicine at the National University of Singapore, and also Co-Head of the World Health Organization Collaborating Centre on Bioethics in Singapore. Calvin has practiced law in London and Singapore with Messrs Linklaters Allen & Gledhill, and served as an Assistant Director of the Legal Aid Bureau (Ministry of Law) in Singapore.

Calvin's research is on global health law and ethics, law and ethics in the context of emerging health technologies (with focus on Artificial Intelligence and data analytics), and on medical law and ethics.

Prof Stephen Li DU
Associate Professor, Faculty of Law, University of Macau, Macau Special Administrative Region, China

Biography:

Prof Li DU is an Associate Professor at the University of Macau, Faculty of Law. Prof Du graduated from Wuhan University, China with dual Bachelor Degrees in both law and clinic medicine in 2005, and received Doctoral Degree in law from the Faculty of Law, University of Alberta, Canada in 2014. His research interests have focused on legal, ethical, and social issues related to cutting-edge biotechnology, e.g., stem cell research and clinical application, genetic testing and gene patents, AI and biomedical research etc.

Mr Zhangyu WANG
PhD Student, Faculty of Law, University of Macau, Macau Special Administrative Region, China

Biography:

Mr Zhangyu WANG is a PhD student at the Faculty of Law, University of Macau. His research interests cover legal issues related to health data protection and cross-border transfer, and the governance of genomics and its interplay with advanced information technologies, including big data and artificial intelligence (AI). He also explores legal regulation issues involving the application of emerging medical, food and feed technologies, such as AI-assisted reproduction technologies, cell-cultured and precision fermentation food and feed.

Dr JI Ping, MD, PhD
Vice Director, Clinical Research Institute, Shenzhen Peking University-Hong Kong University of Science and Technology Medical Center, China

Biography:

Dr JI Ping began her clinical research career as medical writer in 2007 after obtaining her PhD in Cardiff University, UK. She has rich experience in protocol development and project quality management for Phase I to IV new drug development, national and international investigator–initiated studies. Dr Ji also took the role as Director of Quality Assurance Office in Peking University Human Research Protection Program (PU HRPP). Her primary responsibilities were to oversee clinical research projects implementation and assure high quality operation in Peking University Clinical Research Institute between 2010-2016.

Since July 2016, she moved to Shenzhen for building clinical research oversight and support platforms for education & training, technological support, and policy development. Dr Ji has been appointed as Director for Shenzhen Quality Control Center of Clinical Research by Shenzhen Health Committee. As the core member for Shenzhen Biomedical Ethical Review Board, she has made great efforts to develop ethical governance system in biomedical field. In recent years, she has been involved in ethics governance for medical AI, health data, Science and Technology fields in Shenzhen.

Research Interests:
Quality Management in Clinical Research; Governance over Ethics in Biomedical Research

Prof Timo Minssen
Professor of Law
Managing Director and Founder of Center for Advanced Studies and the Collaborative Research Programme in Biomedical Innovation Law (CeBIL)
UNESCO Co-Chair in the Right to Science
University of Copenhagen, Denmark

Biography:

Prof Timo Minssen is Professor of Law and the Founding Director of the Center for Advanced Studies in Bioscience Innovation Law (CeBIL) at the University of Copenhagen. He is also the PI of the NNF's Inter-CeBIL programme, as well as a Research Affiliate at the University of Cambridge's LML and at Harvard Law School's Petrie-Flom Centre for Health Law Policy, Biotechnology & Bioethics. His research, supervision, teaching & advisory practice concentrates on IPRs, Competition & Regulatory Law with a special focus on new technologies, big data & AI in the health & life sciences. Based on his academic and practical experience from law firms, courts & life science start-ups, Timo serves as an advisor and member of expert committees at the WHO, WIPO, EU Commission, various organizations, companies, national governments and law firms. He has published 7 books, as well as 220+ articles and book chapters, which have been featured in i.a. *The Economist, The Financial Times, El Mundo, Politico, WHO Bulletin, Times of India & Times Higher Education*, as well as in leading law journals and science magazines, such as *Science, NEJM Catalyst, JAMA, Harvard Business Review, Harvard Business Manager, Nature Biotechnology, Nature Genetics, Nature Electronics, Nature PJ Digital Medicine & The Lancet Digital Health*.

Dr Colin Mitchell
Head of Humanities, PHG Foundation, University of Cambridge, UK

Biography:

Dr Colin Mitchell is Head of Humanities at the UK based health policy research organisation, the PHG Foundation, which is a linked charity of the University of Cambridge. He leads a team addressing legal, ethical and societal issues that arise with biomedical innovation and new health technologies. This includes challenges associated with genomic data processing, artificial intelligence and the obligations of health and scientific professionals in the contemporary biomedical landscape. Colin has a PhD in health law from the University of Amsterdam, a Masters of Studies in Legal Research from the University of Oxford and a BA in Law from the University of Cambridge.

Prof dr Tom Goffin
Associate Professor of Health Law, Metamedica, Ghent University, Belgium

Biography:

Prof Tom Goffin is a prominent legal scholar and academic known for his expertise in health law, bioethics, and medical AI regulation. Based in Belgium, he serves as an associate professor at Ghent University, where he focuses on the intersection of law, technology, and healthcare. His work often explores the ethical and legal challenges surrounding the use of emerging technologies in medicine, including artificial intelligence (AI), organization of healthcare, and patients' rights. Goffin has contributed significantly to the understanding of how AI can be integrated into healthcare in a trustworthy and legally compliant manner.

In addition to his academic contributions, Goffin plays an active role in advising governmental and international bodies on the legal aspects of healthcare innovation. His research often bridges the gap between theoretical legal analysis and practical regulatory solutions, making him a key figure in shaping the future of health law in Europe. Through his publications, lectures, and consultations, Tom Goffin continues to influence the evolving landscape of medical law, ensuring that technological advancements align with ethical standards legal accountability and healthcare organization.

Prof Oliver Quick
Professor of Health Law and Policy, University of Bristol Law School, UK

Biography:

Prof Oliver Quick is Professor of Health Law and International Director at the University of Bristol Law School. He teaches UG and PG courses on Medical Law, Health Law and Criminal Law (including a Criminal Law course at HKU SPACE). His research is interdisciplinary and impactful and focuses on professionalism, regulation, safety, and trust in healthcare. His monograph *Regulating Patient Safety: the End of Professional Dominance?* (CUP, 2017) was shortlisted for the St Petersburg International Private Law Prize in 2019. Prof Quick has published widely on the need for candour about healthcare harm, comparative systems for incentivising safer maternity care, the criminalization of medical harm, professional fitness to practise frameworks, and the benefits and harms of digital health technologies. He has been a visiting fellow at the University of Auckland, Boston University, University of British Columbia, NUS and UWA.

Prof Sara Gerke
Associate Professor of Law and Richard W. & Marie L. Corman Scholar,
Associate Professor, European Union Center,
University of Illinois Urbana-Champaign, US

Biography:

Prof Sara Gerke is an Associate Professor of Law and Richard W. & Marie L. Corman Scholar at the College of Law, as well as an Associate Professor at the European Union Center, at the University of Illinois Urbana-Champaign. Her current research focuses on the ethical and legal challenges of artificial intelligence and big data for health care and health law in the United States and Europe.

Prof Gerke is leading several research projects funded by the NIH and the European Union. She has over 70 publications in health law and bioethics, especially AI and digital health. Her work has appeared in leading law, medical, scientific, and bioethics journals, including NEJM, JAMA, Science, and Nature Medicine.

Before joining Illinois, Prof Gerke was an Assistant Professor of Law at Penn State Dickinson Law and was promoted early to Associate Professor of Law in 2024. Previously, she served as a Research Fellow in Medicine, Artificial Intelligence, and Law at the Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School, where she oversaw the day-to-day work of the Project on Precision Medicine, Artificial Intelligence, and the Law (PMAIL).

Prof Boris Babic
Associate Professor, HKU Musketeers Foundation Institute of Data Science and Department of Philosophy
Associate Professor (by courtesy), Faculty of Law
The University of Hong Kong, Hong Kong Special Administrative Region, China

Biography:

Prof Boris Babic is Associate Professor of Data Science, Philosophy and (by courtesy) Law at the University of Hong Kong. Previously, he was an assistant professor at the University of Toronto, and an assistant professor at INSEAD (in France and Singapore). He received his PhD in Philosophy and MS in Statistics from the University of Michigan, Ann Arbor, and he received his JD from Harvard Law School. He completed a postdoctoral scholarship at the California Institute of Technology (Caltech) and practiced intellectual property and entertainment law at Quinn Emanuel in Los Angeles. His primary research interests are in legal, ethical, and policy dimensions of artificial intelligence and machine learning, particularly in medical applications.

Dr Barry Solaiman
Assistant Professor of Law, College of Law, Hamad Bin Khalifa University, Qatar, Doha

Biography:

Dr Barry Solaiman is the Associate Dean for Academic Affairs and an Assistant Professor specializing in Healthcare Law at HBKU Law. He is also an Adjunct Assistant Professor of Medical Ethics in Clinical Medicine at Weill Cornell Medicine – Qatar, where he serves as Co-Director of the Intersection of Law and Medicine Conference Series. He holds a PhD in Law from the University of Cambridge and is a Fellow of Harvard Medical School's Center for Bioethics. He was formerly Editor-in-Chief of both the Cambridge International Law Journal and Medicine and Law. He is co-editor of the Research Handbook on Health, AI and the Law, which is the leading book in the field. He has published in leading journals on the regulation of AI in healthcare and was Lead Principal Investigator for a grant at HBKU that created guidelines for the development of AI in healthcare research.

Prof Brian Wong
Assistant Professor, Department of Philosophy, The University of Hong Kong, Hong Kong Special Administrative Region, China

Biography:

Prof Brian Wong is an Assistant Professor in Philosophy at the University of Hong Kong. His research examines the intersection of geopolitics, political and moral philosophy, and technology, with particular interests in the ethics and dynamics of authoritarian regimes and their foreign policies, responding to historical and colonial injustices, and the impact of automation on labour and human societies. Brian is a Fellow at the newly established Centre on Contemporary China and the World, at the University of Hong Kong. As the Chief Strategy Officer of the HK-ASEAN Foundation, he advises multi-national corporations, family offices, and leading think-tanks on geopolitical affairs and macro risks throughout Asia. Having co-founded and served as the inaugural Editor-in-Chief at the *Oxford Political Review*, a publication aspiring to bridge the theory-practice gap, Brian serves as a columnist at the Hong Kong Economic Journal. His writings on Chinese political economy, Asian geopolitics, and public philosophy have been featured in publications such as *TIME, Foreign Policy, Aeon, Financial Times, Diplomat, Fortune, The Hindu, South China Morning Post, Nikkei, Japan Times*, and *the US-Asia Law Institute*. He has also been interviewed by CNN, Al Jazeera, and CGTN for his views on Chinese foreign policy. A Rhodes Scholar (HKSAR, 2020), Brian holds a DPhil in Politics, an MPhil in Political Theory (Distinction), and an MA in Philosophy, Politics, and Economics from the University of Oxford.

Ms Tanya Brigden
Senior Policy Analyst, PHG Foundation, University of Cambridge, UK

Biography:

Ms Tanya Brigden is a Senior Policy Analyst (Biomedical Ethics) at the PHG Foundation, working on ethical and legal considerations arising from biomedical innovation and personalised healthcare. She contributes to a broad portfolio of PHG interests, and has developed expertise in a diverse range of topics including the use of genomic technologies (such as polygenic scores and genome editing), and artificial intelligence in healthcare.

Beyond her role at PHG Foundation, Tanya is the Ethicist member of the NHS Cambridgeshire and Peterborough Clinical Policies Forum, and is a member of the NHS Cambridgeshire and Peterborough System-wide Ethics Committee. Tanya has an MA in Medical Ethics and Law from Kings College London, a Graduate Diploma in Law, and a BA in Philosophy from Durham University

Dr Pete Mills
Director, PHG Foundation, University of Cambridge, UK

Biography:

Dr Pete Mills is the Director of the PHG Foundation, an interdisciplinary health policy research organisation and linked charity of the University of Cambridge, with the mission ‘to make science work for health’. Originally trained in philosophy, Pete has worked for nearly 25 years at the intersection of emerging science, ethics and public policy. Prior to joining the PHG Foundation in 2023, Pete was Associate Director at the Nuffield Council on Bioethics and before that he led the Secretariat of the Human Genetics Commission. He has also held senior policy positions at the UK Department of Health and the Human Fertilisation & Embryology Authority, and has served in representative and advisory roles on several national and international bodies dealing with genomics, bioethics and human rights. For his first degree, Pete read Philosophy, Politics and Economics at Trinity College, Oxford, and he has an MA and PhD in Philosophy from the University of Warwick.

Prof Rachel Sterken
Associate Professor, Department of Philosophy
Co-director of ConceptLab
The University of Hong Kong, Hong Kong Special Administrative Region, China

Biography:

Prof Rachel Sterken is Associate Professor of Philosophy, Chairperson of the Philosophy Department, and Associate Dean (Postgraduate) in the Faculty of Arts at HKU. Prior to joining HKU, she was Associate Professor at the University of Oslo.

Prof Sterken's main research interests are in philosophy of language and communication, conceptual ethics, social epistemology. philosophy and ethics of information/data, and philosophy of technology. Her work is published in leading philosophy journals such as *Philosopher's Imprint*, *Philosophical Studies*, and *Journal of Ethics and Social Philosophy*.

Prof Sterken has taught numerous courses across professional, BA, MA, and PhD levels; including ethics of AI, logic, critical thinking, core ethics, philosophy of language, philosophy of mind, metaphysics, and ethics of information.

Organiser:

CMEL

THE UNIVERSITY OF HONG KONG
CENTRE FOR MEDICAL ETHICS AND LAW

The Centre for Medical Ethics and Law (CMEL) develops new ideas and solutions in response to the big ethical, legal and policy questions of medicine and health. CMEL is the first cross-faculty interdisciplinary institution of its kind in the region. It was founded in 2012 by the LKS Faculty of Medicine and Faculty of Law at The University of Hong Kong as a joint inheritor of their vibrant intellectual traditions dating back to 1887 and 1969 respectively.

Today, CMEL brings together bioethicists, academic lawyers, medical scientists, and other scholars to conduct cutting edge bioethical and legal research and contribute to policy development in flagship areas like digital health and emergent technologies, mental health and capacity, and population and global health.

Research, teaching and knowledge exchange—CMEL's core initiatives—aim to ensure that developments in biomedicine and public health will be underpinned by ethical and legal considerations.

Centre for Medical Ethics and Law, Faculty of Law, The University of Hong Kong

Address:	Office 9.21, 9th Floor, Cheng Yu Tung Tower, Centennial Campus, The University of Hong Kong
Email:	cmel@hku.hk
Tel:	(852) 3917 1845
Fax:	(852) 2549 8495
Website:	https://cmel.hku.hk
Twitter (X):	@HKUCMEL
LinkedIn:	https://www.linkedin.com/company/hku-cmel/
Facebook:	https://www.facebook.com/pg/CMELHKU
Instagram:	https://www.instagram.com/hku_cmel/

Subscribe for updates & newsletters: https://bit.ly/3mCkVaZ

Supporting Organisations:

HONG KONG ACADEMY OF MEDICINE
香港醫學專科學院

HKU Med School of Clinical Medicine
Medical Ethics & Humanities Unit
香港大學醫學倫理及人文學部

Disclaimer:

All the statements and information/ suggestions in this publication are made/ provided without legal responsibility and do not constitute legal or other professional advice. No legal liability shall arise from any errors or omissions in relation to any of the statements/ information/ suggestions in this publication. No express or implied warranties of accuracy or fitness for any particular purpose or use with respect to the statements/ information/ suggestions in this publication are made. The opinions expressed by the contributors in this publication do not necessarily reflect the views of the University of Hong Kong. This publication is for general reference only. The law changes from time to time.

Subscribe cmel.hku.hk

@ HKUCMEL

@ HKU-CMEL

@ CMELHKU

@ HKU_CMEL

CMEL
THE UNIVERSITY OF HONG KONG
CENTRE FOR MEDICAL ETHICS AND LAW