



## Review

# Governance of automated image analysis and artificial intelligence analytics in healthcare



C.W.L. Ho<sup>a,\*</sup>, D. Soon<sup>b</sup>, K. Caals<sup>c</sup>, J. Kapur<sup>d</sup>

<sup>a</sup> Centre for Biomedical Ethics, Yong Loo Lin School of Medicine, National University of Singapore, MD11, 10 Medical Drive, Singapore

<sup>b</sup> Division of Neurology, National University Hospital, Singapore

<sup>c</sup> Department of Geography, Faculty of Arts and Social Sciences, National University of Singapore, Singapore

<sup>d</sup> Department of Diagnostic Imaging, National University Hospital, Singapore

The hype over artificial intelligence (AI) has spawned claims that clinicians (particularly radiologists) will become redundant. It is still moot as to whether AI will replace radiologists in day-to-day clinical practice, but more AI applications are expected to be incorporated into the workflows in the foreseeable future. These applications could produce significant ethical and legal issues in healthcare if they cause abrupt disruptions to its contextual integrity and relational dynamics. Sustaining trust and trustworthiness is a key goal of governance, which is necessary to promote collaboration among all stakeholders and to ensure the responsible development and implementation of AI in radiology and other areas of clinical work. In this paper, the nature of AI governance in biomedicine is discussed along with its limitations. It is argued that radiologists must assume a more active role in propelling medicine into the digital age. In this respect, professional responsibilities include inquiring into the clinical and social value of AI, alleviating deficiencies in technical knowledge in order to facilitate ethical evaluation, supporting the recognition, and removal of biases, engaging the “black box” obstacle, and brokering a new social contract on informational use and security. In essence, a much closer integration of ethics, laws, and good practices is needed to ensure that AI governance achieves its normative goals.

© 2019 The Royal College of Radiologists. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

In April 2018, the US Food and Drug Administration (FDA) granted approval for IDx-DR (DEN180001) to be marketed as the first artificial intelligence (AI)-based diagnostic system that does not require clinician interpretation

to detect greater than a mild level of diabetic retinopathy in adults diagnosed with diabetes.<sup>1</sup> In essence, this ‘Software as a Medical Device’ (SaMD) applies an AI algorithm to analyse images of the eye taken with a retinal camera that are uploaded to a cloud server.<sup>2</sup> A screening decision is made by the device as to whether the individual concerned has “more than mild diabetic retinopathy” and, if so, is referred to an eye care professional for medical attention. Where the screening result is negative, the individual will be rescreened in 12 months. Importantly, IDx-DR is to be applied specifically for detecting diabetic retinopathy, and

\* Guarantor and correspondent: C. W. L. Ho, Centre for Biomedical Ethics, Yong Loo Lin School of Medicine, National University of Singapore, MD11, 10 Medical Drive, 117597, Singapore.

E-mail address: [medhwlc@nus.edu.sg](mailto:medhwlc@nus.edu.sg) (C.W.L. Ho).

in individuals who do not present with higher risks or medical complications.

IDx-DR is capable of machine learning (ML), which is a subset of AI and refers to a set of methods that has the ability to detect patterns in data automatically in order to predict data trends or for decision-making under uncertain conditions.<sup>3</sup> Deep learning (DL) is in turn a subtype of ML (and a subfield of representation learning) that is capable of delivering a higher level of performance, and does not require a human to identify and compute the discriminatory features for it. From the 1980s onwards, DL software has been applied in computer-aided detection systems, and in more recent years, in computer-aided diagnosis (CAD) systems.<sup>4</sup> The field of radiomics (a process that extracts large number of quantitative features from medical images) is broadly concerned with the latter, where DL has enabled the use of CAD systems in computer-learned tumour signatures. It has the potential to detect abnormalities, make differential diagnoses, and generate preliminary radiology reports in the future, but only a few methods are able to manage the wide range of radiological presentations of subtle disease states.<sup>4</sup> Beyond image analysis, AI-based non-image analytical tools may profoundly impact radiology in their potential to improve radiology departmental workflow through precision scheduling, identifying patients who are likely to miss appointments, and producing individually customised examination protocols.<sup>5</sup>

The hype over AI has spawned claims that clinicians like radiologists will become redundant. It is still moot as to whether AI will replace radiologists in day-to-day clinical practice but more AI applications are expected to be incorporated into the workflows of picture archiving and communication systems (PACS) in the foreseeable future. Areas that can be automated with AI have been identified as<sup>6,7</sup>: (a) automated image segmentation, lesion detection, measurement, labelling and comparison with historical images; (b) generating radiology reports, particularly with the application of natural language processing and natural language generation; (c) semantic error detection in reports; (d) data mining in research; and (e) improved business intelligence systems that allow real-time dash-boarding and alert systems, workflow analysis and improvement, outcomes measures and performance assessment. Policy and professional interventions may become necessary to manage job displacement as repetitive, low discretion, and mundane tasks become automated; to create new jobs and roles within healthcare (such as medical data scientists); and to reduce friction from the transition.

It is consequently likely that AI applications will produce significant ethical and legal issues in healthcare, especially if they cause abrupt disruptions to its contextual integrity and relational dynamics. Healthcare is currently delivered by licensed professionals at accredited facilities or licensed premises, and designed to flow from a professional through approved channels to care recipients. The professional here is also the repository of knowledge, which underscores the notion of a fiduciary relationship where a clinician is required to act in the best interests of her or his patient.<sup>8</sup> AI applications are part of a rapidly growing number of new

technologies that directly or indirectly enable lay people to access a rich pool of knowledge, interact with other individuals with expertise and/or experiential knowledge and may eventually be able to derive accurate diagnoses and develop effective healthcare regimens independently of a clinician. AI applications could also disrupt current healthcare practices. In many health systems, clinical medicine increasingly reflects a shift-based model, where fewer clinicians follow diseases from their presentation through their ultimate outcome.<sup>9</sup> AI applications may assume a greater role in integrating healthcare, even while a clinician remains at the centre in attribution of ethical and legal responsibilities, in payment arrangements, and in organisational setup. In addition, AI may curtail a patient's exercise of her or his right to privacy and confidentiality as ML analysis will require the patient's personal information to be recorded.<sup>9</sup>

Disruptions need not necessarily be negative, provided that the goals of medicine are sustained. In this paper, we focus on one goal in particular: service of the sick or otherwise acting for the good of patients. It follows that our use of the term "clinician" is intended to refer broadly to a moral community of professions (especially nurses and doctors) that is socially established and trained to serve patients.<sup>10</sup> In this regard, we are concerned about the impact of AI applications on trust on the part of patients, as well as trustworthiness on the part of clinicians and health systems. Sustaining trust and trustworthiness in healthcare is not simply a professional concern.<sup>11</sup> It is recognised in ethics and in law to be in the public interest, and consequently, it is also a key goal of healthcare governance. In the section that follows, we consider the nature of governance in AI-based biomedicine and the limits to achieving its normative goals. We then consider how concerted professional responses are needed, particularly on the part of the radiology community.

## General trend in governance

As the discussion in this section seeks to illustrate, the nature of governance of biomedicine is increasingly risk-based, context-specific, case-sensitive, decentralised, collaborative and, in terms of its epistemic (knowledge) constituents, pluralistic. As we considered above, IDx-DR was reviewed under the FDA's De Novo premarket review pathway and was granted Breakthrough Device designation,<sup>12</sup> as the SaMD is novel and of low to moderate risk. Importantly, the FDA provided intensive interaction and guidance to the company on efficient device development in order to expedite evidence generation and its review of the device that provides for more effective treatment or diagnosis of a life-threatening or irreversibly debilitating disease or condition, with no approved or cleared alternatives.

A risk-based approach is adopted by the FDA, and its application to a total product life-cycle process is embodied in ISO 14971, as well as its accompanying guidance document, ISO TR 24971 (an ISO Technical Report).<sup>12,13</sup> These documents outline a risk management process for medical

device manufacturers to identify hazards (defined as potential sources of harm), estimate and evaluate risks, and to respond effectively through developing, implementing, and monitoring risk control measures. Broadly speaking, risk analysis performed within this framework entails the systematic use of available information to identify hazards and to estimate the risk that arises from the scope and intended use of the device. Risk estimates are computed based on the probability of occurrence of harm and the severity of that harm. These estimates are then evaluated, such as through the use of a risk evaluation matrix, which distinguishes acceptable levels of risks from unacceptable ones. Risk controls (through product design or labelling) are developed and implemented to reduce unacceptable risks. The effectiveness of each risk control measure must be evaluated to ensure that residual risks are within acceptable levels. Otherwise, additional risk control measures will need to be introduced. Additionally, a similar analysis must be conducted to evaluate the entire device (that is, analysis that is not linked to particular hazard) in order to determine if the overall residual risk meets that acceptability criteria of the manufacturer or developer. A risk management report documents the entire process, and is updated with post-development risk-relevant events such as product feedback and non-conformance. This standard was last reviewed and confirmed in 2010, but is currently being reviewed.<sup>14</sup>

Similar to the approach in the US, the European regulatory regime<sup>15–17</sup> adopts a risk-based approach to regulate medical devices, which are defined as any instrument or tool (including software) intended by the manufacturer to be used for humans for purposes that include diagnosis, prevention, monitoring, treatment, or alleviation of disease.<sup>18</sup> The responsibility of risk assessment is placed on device manufacturers or on an independent certification body appointed by authorities of EU member states. Recent changes to the regulatory framework will mostly come into effect from 2020 onwards,<sup>19,20</sup> and they will widen the range of products that will be regulated, extend liability to defective products, strengthen requirements for clinical data and traceability of the devices, introduce more rigorous monitoring of certification bodies and improve transparency by requiring more product information to be placed in the public domain.<sup>21</sup> Such a governance approach reflects what Matthew Scherer<sup>22</sup> terms an *ex ante* (as opposed to an *ex post*) regulation. It is pre-emptive of foreseeable risks and has a more open and participatory character. Although this approach has been effective in paving a way forward to market approval for IDx-DR, the study that supports the approval<sup>23</sup> was conducted under highly controlled conditions where a relatively small group of carefully selected patients was recruited to test a diagnostic system under narrow usage criteria. It is questionable whether the AI feature itself was tested, as the auto-didactic aspect of the algorithm was locked prior to the clinical trial, which greatly reduced the range of possible outputs.<sup>24</sup> At this stage, IDx-DR is not capable of evaluating the most severe forms of diabetic retinopathy that require urgent ophthalmic intervention. Moving forward, unsupervised

ML devices will test the limits of this governance approach. The challenges to risk assessment, management, and mitigation will be amplified as AI-based devices change rapidly, with the assumption of less tangible form (discreteness), embodiment of more diverse components (discreteness), greater dispersion across geographical and jurisdictional spaces (diffuseness) and opacity.<sup>22</sup>

Governance is likely to become increasingly complex, less certain, and more challenging to describe. In the infosphere, Floridi<sup>25</sup> depicts governance (along with regulation and ethics) as one of the three normative cornerstones. Digital governance is the practice of establishing and implementing policies, procedures, and standards for the proper development, use, and management of the infosphere. It overlaps with (but is not identical to) digital regulation, which relates to a system of laws elaborated and enforced through social and governmental institutions to regulate behaviour in the infosphere. The General Data Protection Regulation (GDPR) that came into force last year in Europe is an example of digital regulation.<sup>26,27</sup> Governance and regulation may be shaped by digital ethics, which is concerned with moral problems relating to data and information, algorithms, and corresponding practices and infrastructures. One may perhaps query if ethics and regulation are as discrete as Floridi conceptualises them, but for the purposes of this paper, ethics and regulation are regarded as constitutive of, and indistinguishable from, governance.<sup>28</sup>

Governance, in the form of bioethics, has been central to the advancement of biomedicine, as empirical studies have shown.<sup>29–31</sup> Where “big data” and AI-based biomedicine is concerned, sound and effective governance remains crucial to all endeavours that are directed at understanding and responding to the many challenges that are posed and anticipated. This is evident in the policy documents that have been produced at the highest level of government on what good governance could and perhaps should look like. In the US, the Executive Office of the President<sup>32,33</sup> considered how good governance of AI could advance national priorities, broadly enumerated as increased economic prosperity, improved educational opportunity, social security, and quality of life, and enhanced national security. Public good, fairness, and safety have been identified as key guiding principles. The US National Science and Technology Council<sup>34</sup> provided a more nuanced roadmap on how some of these priorities may be realised through investments in research and development. In the UK, a report of the House of Commons’ Science and Technology Committee on AI<sup>35</sup> is similar to the US policy documents on AI in its call for a “light touch” approach to governance for sustaining technological development to advance national economic interests. Of especial note is its call for the UK government to provide leadership on the ethical, legal, and social implications (ELSI) of AI, even if it should trail behind other countries such as the US, Germany, and Japan on the technical aspects. Across the English Channel, the European Parliament’s Committee on Legal Affairs<sup>36</sup> proposed for a regulatory agency to be established in order to address ethical and legal issues concerning AI-enhanced robots.

Arguably, these are still early days in thinking about desirable (let alone ideal) governance of AI-based biomedicine in terms of its normative and conceptual foundations, mechanisms, and practices. Not surprisingly, our colleagues<sup>37</sup> observe that the vision of the “good AI society” remains obscure, even if we can generally agree that such a society should be animated by respect for human dignity and human rights, rule of law, democracy, transparency, and accountability. Current “real world” arrangements present difficult questions as to how we should think about values like democracy, transparency, and accountability when AI technology appears to be primarily driven by large commercial interests. Many more colleagues have highlighted concerns over the unchecked use of AI systems in finance, education, criminal justice, search engines and social welfare that not only lacks normative justification, but can give rise to serious detrimental effects.<sup>38</sup> Related concerns about fairness and equity, privacy and security, and trust are well articulated in important ethical and scientific forums,<sup>39–41</sup> even as uncertain legal and social conditions (over intellectual property protection for AI and AI-derived work, for instance) greatly obfuscate normative evaluation. Concerns over the disproportionate influence of specific interests over AI governance cannot be understated, especially if “light touch” approach is to carry the day. Important questions proffered in the literature include: Who sets the agenda for AI governance? What values and cultural logic are instantiated? Who ultimately benefits from it? In “big data” and AI-based biomedicine, clinicians in general and radiologists in particular will have greater ethical, legal, and social responsibilities in ensuring that AI-related technologies are appropriately understood, developed, and applied.

## Professional responsibilities

Public hype and professional anxiety over potential job losses as a consequent of AI could be a red herring to the need for more measured and consultative professional leadership in the governance of this emergent technology. Securing trust and maximising value for patients and health systems should remain central concerns of all clinicians, with or without AI intervention, and also the goals of good AI governance. Although AI tools are often presented as socially desirable because they make healthcare more efficient and affordable,<sup>42</sup> their clinical and social value may not always be well established through inquiry and consensus. Many research and clinical applications in radiology including molecular imaging, radiometric, radiogenomics, and large population cancer screening will involve data mining, which will be facilitated by ML analyses.<sup>7</sup> In order to advance AI-based imaging research and practice, supportive infrastructure and operational standards must be in place to integrate AI applications and ensure that they are interoperable with existing clinical radiology workflows, based on their role, type, and use cases. These include establishment of national and international image sharing networks, acquisition protocols,

reference datasets of proven cases against which AI-based software can be tested and compared, criteria for standardisation, validation and optimisation of imaging protocols for use in AI applications, and a common lexicon for describing and reporting these applications.<sup>43</sup> Robust methods need to be developed for quality control of shared images and for ensuring the integrity of image data. Standards for curation of images will also need to be established. It is unclear if any country has established a comprehensive informatics system capable of supporting data transfer, storage, quality control, and management across healthcare and research institutions, but high variability in imaging protocols among institutions is a key obstacle in many health systems. In addition, medical images are highly heterogeneous at both an individual and a population level. These issues limit the dataset that can be composed for training and validation purposes, and create a risk of “overfitting” the data and loss of generalisability, even if there are different techniques to identify and reduce overfitting.<sup>44,45</sup> In the light of significant differences in disease prevalence, imaging protocols with different imaging characteristics, choice of reference standard and equipment both within and across different countries, the scope of application of an AI software will need to be critically evaluated.<sup>5,6</sup> Upscaling AI-based techniques will require a much clearer understanding of the clinical need (or use case) and the business case (if commercialised), product regulation, verification, and monitoring.

Organisational changes, encompassing definition of roles, technical considerations, and requirements for implementation, are also needed. For instance, integration of CAD with PACS will allow for clinical routine use of CAD as an adjunct to radiologist interpretation of all modalities. Additionally, integration with electronic health records and radiology information systems will enable access to patients’ supporting information, history, test results, and images. As Liew<sup>7</sup> explains, the stakeholders needed to bring about these changes include the Chief Information Officer (CIO) and, if available, Chief Data Officer (CDO) of each hospital, radiology leadership in committees and academic bodies such as professional colleges and societies, as well as individual radiologists. The CIO’s role is to ensure that AI-based initiatives can be implemented safely and effectively, maintain integrity, and (where appropriate) transferability of data in the electronic health data systems and consistency with the hospital’s data policies. The CDO in some organisations are seen to have a role that is supportive of the CIO, and is responsible for ensuring data quality for validation and training of ML systems and compliance with governance requirements. Their work must not be viewed narrowly only as purely institutional good practices, but as professional concerns that should be supported by professional organisations. In the foreseeable future, clinicians (and especially radiologists) will be critical gatekeepers and intermediators to AI-based imaging devices and systems. As AI in biomedicine is not self-enabling, professional standards relating to research and practice will become even more important, even as product or device regulation gain in complexity.



### *Inquire into clinical and social value*

A gulf is said to exist between an algorithm that works well on a small dataset from a specific population and one that can be reliably applied to a large population and across different imaging methods.<sup>24</sup> It is still unclear when a proof-of-concept research can be adopted and upscaled as a marketed product. At least one study suggests that AI-based CAD has not improved the diagnostic accuracy of mammography and may even have resulted in missed cancer diagnoses. Consequently, insurers could have paid more for CAD with no established benefit to patients.<sup>46</sup> Studies in the UK and the US on AI-based CAD for breast cancer show a significant improvement in the cancer detection rate in comparison to analysis by an individual radiologist only, along with an acceptable increase in recall rate and minimal impact on positive predictive value<sup>47,48</sup>; however, when compared with an arrangement whereby a second radiologist reviews the image in lieu of the AI-based CAD, the evidence of benefit is more ambiguous. Even if the CAD is more effective in picking up features that may be of concern, it still remains up to the initial reader to determine whether medical intervention is necessary. If the initial reader did not pick up a suspected feature owing to bias, then (s)he will probably still dismiss it when it is highlighted by the CAD. In contrast, if a second human reader disagrees with the initial reader, then that second reader is more likely to ensure appropriate medical follow-up.<sup>49–51</sup> The economic value of AI-based CAD in imaging is similarly unclear, as data on cost effectiveness is generally limited. The use of such software for breast cancer screening in the UK was not found to be cost effective when compared with double-reading unless recall rate could be significantly reduced, for instance.<sup>52</sup>

Our intention here is not to pronounce on the clinical and social value of AI-based CAD. Instead, we are of the view that a more comprehensive dialogue that involves a broad range of interested parties (including lay members of the public) is urgently needed. Professional bodies have a special role to play in initiating and moderating this dialogue, in view of their domain expertise, social standing, and public responsibilities. Most ML capability for image analysis in radiology is performed via supervised learning which requires appropriately labelled training data. Adequate labelling of key imaging findings is a tedious and time-consuming process, and is further dependent on clear demarcation of normal and abnormal features.<sup>5</sup> Additionally, the investments needed to develop AI-based software are substantial. It is unclear how institutions can support AI core laboratories while also allowing access to others. Infrastructural improvements are needed within a health-care facility, which could involve computational hardware procurement and upgrades, and ensuring connectivity to secure cloud platforms and data storage. It may also be necessary to organise additional training for IT support staff and staff in general, not only on the technical aspects but also on the organisational ones, such as data protection and cyber-security practices. Cost constraints may require the

prioritisation of some AI capabilities over others, and any such determination should satisfy ethical and legal requirements.<sup>53</sup>

### *Alleviate deficiencies in technical knowledge*

A number of radiology associations and committees from professional colleges and societies already recognise the need to provide guidance and set clear standards for the entire professional body and interested stakeholders. These guidance and standards could relate to advancing and implementing AI software in radiology, validating AI software, and developing a general implementation roadmap for the future. Conceivably, some of these tasks should be undertaken by professional bodies collaboratively with regulators and policymakers. Currently, the American College of Radiology has provided some guidance in defining use cases for common problems,<sup>54</sup> while the Royal College of Radiologists has issued a position statement on AI and has established an AI Framework.<sup>55,56</sup> A more comprehensive set of recommendations has been published by the AI Working Group of the Canadian Association of Radiologists (CAR-WG), including those that relate to professional governance.

The CAR-WG highlights the need for radiologists to be familiar with different AI techniques, to understand the challenges related to the preparation of training datasets for supervised learning, and to be familiar with AI terminology and hierarchy. The radiology community needs to be educated on how to critically analyse the opportunities, pitfalls, and challenges associated with the introduction of new AI tools.<sup>5,9</sup> For instance, false-positive rates continue to be a pitfall for CAD software. Although false-positive rates have been reduced over time through redesigning and retraining, the tendency for such software to incorrectly identify normal structures as abnormal remains one of the key concerns. Substantial time and resources are needed to determine if the highlighted concerns are true abnormal lesions, often negatively impacting recall rates, quality of healthcare service delivery and cost.<sup>51</sup> Guidance should also be developed on appropriate means of communicating effectively with patients, along with supportive counselling, where applicable.

### *Support recognition and removal of biases*

Radiologists and clinicians using AI techniques must appreciate the values that may either be built into the system or introduced through the data used in ML algorithms. AI devices and systems may be trained to support practices that reflect some ethical goals but not others, and may even be trained to circumvent legitimate ethical and legal requirements. ML models can also incorporate implicit selection biases from the demographics of the population used for its training, which may not be representative of the target population in which it will be applied.<sup>5</sup> Algorithms used in non-medical fields show how they mirror human biases in decision making. Famously, racial discrimination

was reflected by programs designed to aid judges in sentencing because the risk of recidivism has been presented to be disproportionately higher for certain racial groups.<sup>57</sup> Within a healthcare setting, if clinicians always withdraw care in patients with a brain injury, a ML system may conclude that this action should always follow from such a finding.<sup>9</sup> Other ethical and legal contentions may arise from differences between the intended use of the AI device or system and the goals of the users. As the designer and purchaser of a ML system are usually not those who deliver bedside care, values and priorities that are embedded and entailed may very well diverge significantly.<sup>9</sup> How should ethical dilemmas be dealt with if there is no consensus among experts or interested stakeholders? More guidance will eventually be needed on the normative content and skills that should be part of clinical education relating to biases in AI applications.

### *Engaging the “black box” obstacle*

A well-known concern with AI-based CAD is that it is unable to explain its decision. Due to its reliance on a very large number of single associations, it is very difficult to identify the technical and logical reasoning behind any decision made by ML software. It is also currently difficult, if not impossible, to say if such a software is more likely or not to detect a rare disease condition.<sup>4,6</sup> In the event of a misdiagnosis due to an error attributable to the CAD software, there may be further complication as to whether the manufacturer or the radiologist should be responsible.<sup>58</sup> If the legal standard of care is defined by professional practice standards, a radiologist (or indeed any clinician) should be wary of relying on a recommendation from the CAD software that cannot be explained. Professional bodies will need to engage with regulatory agencies on this and other medico-legal issues, and develop common standards for evaluating, validation, and testing AI tools.<sup>5</sup>

### *Brokering a new social contract on informational use and security*

The obligation to protect the confidentiality of health information has long been recognised to serve the public interest of encouraging free and open communication in a trusting doctor–patient relationship. Today, the obligation of confidentiality has been codified for both medical practice and research in regulation and in professional guidelines. Its operative rationale is the ethical imperative of respect for the dignity of persons as autonomous agents. Some countries have introduced legislation to address concerns relating to particular informational risks. This has been the approach in the US where there is no constitutional right to informational privacy. Three US federal legislations provide different forms of control over certain kinds of health information: Privacy Act,<sup>59</sup> Health Insurance Portability and Accountability Act (HIPAA),<sup>60</sup> and Genetic Information Nondiscrimination Act (GINA).<sup>61</sup> The Privacy Act prevents unauthorised disclosure of personal information held by the US federal government, and the persons to

whom the information relates (i.e., data subjects are conferred with certain rights in relation to its accessing, processing, and maintenance). HIPAA protects potential and current employees from discrimination by health insurers and employers, but it has been criticised as a disclosure regulation rather than a privacy rule, as it permits broad and easy dissemination of patients’ medical information, with no audit trails for most disclosures.<sup>62</sup> GINA expands the protection against certain discriminatory practices under HIPAA, as updated and revised by the Health Information Technology for Economic and Clinical Health Act,<sup>63</sup> and it provides a level of protection against genetic discrimination by disallowing health insurers and employers to use certain types of genetic information, but it does not address the use of or access to genetic data, or otherwise provide comprehensive privacy protections. As for US state laws, there is similarly no standard or comprehensive approach to the protection of genetic information, with the level of protection varying widely from state to state.<sup>64</sup> Where cybersecurity is concerned, HIPAA requires healthcare institutions to protect their systems and the personal data that is under their control; however, it does not specify what cybersecurity measures must be implemented. For medical devices, the FDA (working collaboratively with the US Department of Homeland Security) requires manufacturers and healthcare delivery organisations to report a limited number of risks presented by their devices and to take steps in ensuring that appropriate safeguards are in place. This may be challenging for AI-based systems if unsupervised ML capability is put into effect, although the FDA<sup>65</sup> has announced further regulatory developments and activities that are expected to follow.

In Europe, the law on informational privacy is set out in the GDPR, which replaces the 1995 Data Protection Directive.<sup>66</sup> Having come into force in May 2018, the GDPR not only extends the territorial scope of data protection for patients, research participants, and data subjects in general, but also widens the rights that are intended to secure the autonomy, dignity, and privacy of these individuals. It gives considerable emphasis to the requirement of explicit informed consent for handling of health information, even if exceptions (such as processing for research) are applicable, and also confers on the data subjects several other rights, such as a right to information about how the personal data are being processed, a right to rectification of the data, and right to a judicial remedy for any breach of these rights; however, the application of data protection law on medical research (including those that relate to AI-based applications) is not always clear and straightforward, particularly if oversight bodies are outside of the EU. Where cybersecurity is concerned, a new directive has been implemented in May 2018, where EU member states are required to prevent cyberattacks, mitigate any harms that arise as a consequent, and maintain the provision of essential services.

Health information that has been sufficiently anonymised will not fall within the purview of data protection regimes. Where the development of AI-based software is concerned, sensitive information may be derived illicitly

from unknown sources and in respect of which there may not be clear regulatory or legal requirements or oversight.<sup>21,67,68</sup> Large datasets that are representative of the scope and variety of patient types and disease conditions, and available in machine-readable form, are needed to train and test AI algorithms, along with complete and well-structured metadata (including information on data sources). To exacerbate the challenge of data shortage, the validation dataset should ideally not overlap with the testing dataset in order to ensure that the algorithm delivers reliable and accurate output under a wide-range of conditions that could realistically be encountered in a real-world setting within which it is applied.<sup>69</sup> Here again, professional bodies are well-placed to broker a social contract involving policymakers, interested (non-governmental) stakeholders and the general public on the appropriate forms of social control over personal information, the extent that informational resources may be applied in the public interest, appropriate safeguards that should be in place and benefit sharing.

## Conclusion

AI-based CAD software could eventually help radiologists to better manage workload, enhance individual performance and reduce human error. Although this software device is currently limited to a single common disease, its capability is expected to grow over time to identify multiple challenging diagnoses.<sup>45</sup> In the foreseeable future, other AI-based devices and systems will also be assimilated into clinical practice with a view to improve quality and efficiency, similar to current applications on digital imaging methods.<sup>70</sup> These changes could shift the current practice of placing patients within risk groups and move it towards more individualised predictions; however, the radiology community must first assume a more active role in propelling medicine into the digital age. This could certainly be done by combining AI and radiologists into a form of hybrid intelligence or a “radiologist-in-the-loop”.<sup>7</sup> At a professional level, clinicians will need to work closely with the AI research and development community, regulators and policymakers to realise the vision of a dynamic, adaptive, and active learning healthcare system.

More immediate concerns include encouraging the development of representative training datasets, adoption of a common interoperable software framework for research and clinical purposes, and introduction of a standardised approach for benchmarking and implementation of AI applications.<sup>5</sup> As we have discussed, the success of AI applications will depend on researchers gaining access to large volume of quality health data. This will certainly be the case for upscaling IDx-DR, and the road ahead for similar software will be challenging. A higher regulatory threshold under the FDA’s De Novo premarket review pathway now applies, as a diabetic retinopathy AI diagnostic device is already available in the market. Fundamental to all of these initiatives is the need to develop and sustain a governance approach that promotes

collaboration among all stakeholders to ensure the responsible development and implementation of AI in biomedicine. Such an approach will require a much closer integration of ethics, laws, and good practices,<sup>71</sup> and is viable only if it engenders trust and is trustworthy.

## Conflict of interest

The authors declare no conflict of interest.

## References

- Food & Drug Administration. *FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems*. 11 April 2018 <https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm604357.htm>. Accessed 5 March 2019.
- Food & Drug Administration. *Is the product a medical device?*. US Department of Health and Human Services; 2018 <https://www.fda.gov/medicaldevices/deviceregulationandguidance/overview/classifyyourdevice/ucm051512.htm>. Accessed 5 March 2019.
- Murphy KP. *Machine learning: a probabilistic perspective*. Cambridge MA: MIT Press; 2012.
- Giger ML. Machine learning in medical imaging. *J Am Coll Radiol* 2018;**15**:512–20.
- Tang A, Tam R, Cadrin-Chênevert A, et al. Canadian association of radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J* 2018;**69**(2):120–35.
- Lee JG, Jun S, Cho YW, et al. Deep learning in medical imaging: general overview. *Korean J Radiol* 2017;**18**(4):570–84.
- Liew C. The future of radiology augmented with artificial intelligence: a strategy for success. *Eur J Radiol* 2018;**102**:152–6.
- Richman B. Health regulation for the digital age – correcting the mismatch. *N Engl J Med* 2018;**379**(18):1694–5.
- Char DS, Shah NH, Magnus D. Implementing machine learning in health care – addressing ethical challenges. *N Engl J Med* 2018;**378**(11):981–3.
- Sulmasy DP. Edmund Pellegrino’s philosophy and ethics of medicine: an overview. *Kennedy Inst Ethics J* 2014;**24**(2):105–12.
- O’Neill O. *Autonomy and trust in bioethics*. Cambridge: Cambridge University Press; 2002.
- Food & Drug Administration. *Software as a medical device (SAMd): clinical evaluation*. US Department of Health and Human Services; 2017 <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM524904.pdf>. Accessed 5 March 2019.
- International Organization for Standardization (ISO). *Draft International Standard ISO/DIS 14971: medical devices – application of risk management to medical devices*. 2018 [https://s3-eu-west-1.amazonaws.com/static.wm3.se/sites/16/media/227639\\_prEN\\_ISO\\_14971.pdf?1535906215](https://s3-eu-west-1.amazonaws.com/static.wm3.se/sites/16/media/227639_prEN_ISO_14971.pdf?1535906215). Accessed 5 March 2019.
- Bills E. *A Look at the ISO 14971 and ISO TR 24971 Updates*. *Med Device Online*. 15 August 2018 <https://www.meddeviceonline.com/doc/a-look-at-the-iso-and-iso-tr-updates-0001>. Accessed 5 March 2019.
- European Economic Community. *Council Directive on the approximation of the laws of the Member States relating to active implantable medical devices*. 90/385/EEC. Official Journal of the European Communities; 1990.
- European Economic Community. *Council directive concerning medical devices*. 93/42/EEC. Official Journal of the European Communities; 1993.
- European Commission. *Directive 98/79/EC of the European Parliament and the Council on in vitro diagnostic medical devices*. Official Journal of the European Communities; 1998.
- European Commission. *MDCG 2018-2 Future EU medical device nomenclature – description of requirements*. 2018.
- European Parliament and the Council of the European Union. *Regulation (EU) 2017/745 of the European parliament and of the Council 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*. Official Journal of the European Communities; 2017.
- European Parliament and the Council of the European Union. *Regulation (EU) 2017/746 of the European parliament and of the Council on in vitro*



- diagnostic medical devices and repealing directive 98/79/EC and commission decision 2010/227/EU. Official Journal of the European Communities; 2017.
21. Pesapane F, Volonte C, Codari M, et al. Artificial intelligence as a medical device in Radiology: ethical and regulatory issues in Europe and the United States. *Insights Imaging* 2018;**9**(5):745–53.
  22. Scherer MU. Regulating artificial intelligence systems: risks, challenges, competencies and strategies. *Harv JL Tech* 2016;**29**:354–400.
  23. Abramoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *Digital Med* 2018;**39**(1):1.
  24. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. *Digital Med* 2018;**1**:40.
  25. Floridi L. Soft ethics, the governance of the digital and the general data protection regulation. *Phil Trans R Soc A* 2018;**376**:20180081.
  26. European Parliament and the Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.
  27. Marelli L, Testa G. Scrutinizing the EU general data protection regulation. *Science* 2018;**360**(6388):496–8.
  28. Laurie G. What does it mean to take an ethics+ approach to Global Biobank Governance. *Asian Bioeth Rev* 2017;**9**(4):285–300.
  29. Jasanoff S. *Designs on nature: science and democracy in Europe and the United States*. Princeton: Princeton University Press; 2005.
  30. Nowotny H, Testa G. *Naked genes: reinventing the human in the molecular age*. Cambridge MA: MIT Press; 2010.
  31. Ho CWL. *Juridification in bioethics: governance of human pluripotent cell research*. London: Imperial College Press; 2016.
  32. Executive Office of the President. *Artificial intelligence, automation and the economy*. December 2016. Washington DC, <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF>. Accessed 5 March 2019.
  33. Executive Office of the President/National Science and Technology Council, Committee on Technology. *Preparing for the future of artificial intelligence*. October 2016. Washington DC, [https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf). Accessed 5 March 2019.
  34. Executive Office of the President, National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee. *The national artificial intelligence research and development strategic plan*. Washington DC: US Government; October 2016 [https://www.nitrd.gov/PUBS/national\\_ai\\_rd\\_strategic\\_plan.pdf](https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf). Accessed 5 March 2019.
  35. House of Commons Science and Technology Committee. *Robotics and artificial intelligence: fifth report of session 2016–17. HC 145*. 12 October 2016. London, <https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf>. Accessed 5 March 2019.
  36. European Parliament Committee on Legal Affairs. *European civil law rules in robotics: study for the JURI committee*. Brussels: European Parliament; 2016 [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/5713\\_79/IPOL\\_STU\(2016\)571379\\_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/5713_79/IPOL_STU(2016)571379_EN.pdf). Accessed 5 March 2019.
  37. Cath C, Wachter S, Mittelstadt B, et al. Artificial intelligence and the 'good society': the US, EU, and UK approach. *Sci Eng Ethics* 2018;**24**:505–28.
  38. Cath C. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philos Trans Royal Soc A* 2018;**2018**(2133):20180080.
  39. AI100 Standing Committee and Study. *Artificial intelligence and life in 2030*. Stanford University; 2016 <https://ai100.stanford.edu/2016-report>. Accessed 5 March 2019.
  40. Nuffield Council on Bioethics. *Artificial intelligence (AI) in healthcare and research*. London: Nuffield Council on Bioethics; 2018 <http://nuffieldbioethics.org/project/briefing-notes/artificial-intelligence-ai-healthcare-research>. Accessed 5 March 2019.
  41. National Academy of Sciences. Artificial intelligence and machine learning to accelerate translational research. In: *Proceedings of a workshop in brief*. Washington DC: National Academy of Sciences; 2018. <https://www.nap.edu/catalog/25197/artificial-intelligence-and-machine-learning-to-accelerate-translational-research-proceedings> Accessed 5 March 2019.
  42. Bernaert A, Akpakwu E. *Four ways AI can make healthcare more efficient and affordable*. World Economic Forum; 31 May 2018 <https://www.weforum.org/agenda/2018/05/four-ways-ai-is-bringing-down-the-cost-of-healthcare/>. Accessed 5 March 2019.
  43. Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *J Am Coll Radiol* 2018;**15**(3 Pt B):504–8.
  44. Ravi D, Wong C, Deligianni F, et al. Deep learning for health informatics. *IEEE J Biomed Health Inform* 2017;**21**(1):4–21.
  45. Erickson BJ, Korfiatis P, Kline TL, et al. Deep learning in radiology: does one size fit all? *J Am Coll Radiol* 2018;**15**:521–6.
  46. Lehman CD. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;**175**:1828–37.
  47. Sahiner B, Chan HP, Hadjiiski LM, et al. Effects of CAD on radiologists' detection of lung nodules on thoracic CT scans: analysis of an observer performance study by nodule size. *Acad Radiol* 2009;**16**(12):1518–30.
  48. Liang M, Tang W, Xu DM, et al. Low-dose CT screening for lung cancer: computer-aided detection of missed lung cancers. *Radiology* 2016;**281**(1):279–88.
  49. Georgian-Smith D, Moore RH, Halpern E, et al. Blinded comparison of computer-aided detection with human second reading in screening mammography. *Am J Roentgenol* 2007;**189**(5):1135–41.
  50. Azavedo E, Zackrisson S, Mejare I, et al. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. *BMC Med Imaging* 2012;**12**:22.
  51. Fazal MI, Patel ME, Tye J, et al. The past, present and future role of artificial intelligence in imaging. *Eur J Radiol* 2018;**105**:246–50.
  52. Guerriero C, Gillan MG, Cairns J, et al. Is computer aided detection (CAD) cost effective in screening mammography? A model based on the CA-DET II study. *BMC Health Serv Res* 2011;**11**:11.
  53. Daniels N, Sabin JE. *Setting limits fairly*. New York: Oxford University Press; 2002.
  54. American College of Radiology. TOUCH-AI directory. Available at: <https://www.acrdsi.org/DSI-Services/TOUCH-AI>. Accessed 5 March 2019
  55. Royal College of Radiologists. *RCR position statement on artificial intelligence*. 20 July 2018. Available at: <https://www.rcr.ac.uk/posts/rcr-position-statement-artificial-intelligence>. Accessed 5 March 2019.
  56. Royal College of Radiologists. AI: our framework. Available at: <https://www.rcr.ac.uk/policy-public-and-media/policy/artificial-intelligence/ai-our-framework>. Accessed 5 March 2019
  57. Angwin J, Larson J, Mattu S, et al. *Machine bias: there's software used across the country to predict future criminals. And it's biased against blacks*. ProPublica; 23 May 2016 <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed 5 March 2019.
  58. Pinto A, Brunese L, Pinto F, et al. The concept of error and malpractice in radiology. *Semin Ultrasound CT MR* 2012;**33**(4):275–9.
  59. *US congress. privacy act, 5 USC § 552a*. 1974.
  60. Congress US. Health insurance portability and accountability act. *Public Law* 1996;**104–191**.
  61. Congress US. Genetic information nondiscrimination act. *Public Law* 2008;**110–233**.
  62. Sobel R. The HIPAA paradox: the privacy rule that's not. *Hastings Cent Rep* 2007;**37**(4):40–50.
  63. Congress US. Health information technology for economic and clinical health act. *Public Law* 2009;**111–9**.
  64. US Presidential Commission for the Study of Bioethical Issues. *Privacy and progress in whole genome sequencing*. Washington DC: Presidential Commission for the Study of Bioethical Issues; 2012.
  65. Food & Drug Administration. *Cybersecurity*. US Department of Health and Human Services; 2018 <https://www.fda.gov/medicaldevices/digitalhealth/ucm373213.htm>. Accessed 5 March 2019.
  66. European Parliament and the Council of the European Union. *Directive 95/46/EC of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Official Journal of the European Communities; 1995.
  67. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep* 2014;**16**:441.



68. Castelvechi D. Can we open the black box of AI? *Nature* 2016;**538**:20–3.
69. Lyapustina S. US FDA approaches to artificial intelligence. *The national law review*. 24 April 2018 <https://www.natlawreview.com/article/us-fda-approaches-to-artificial-intelligence>. Accessed 5 March 2019.
70. Yeung S, Downing L, Li FF, et al. Bedside computer vision — moving artificial intelligence from driver assistance to patient safety. *N Engl J Med* 2018;**378**(14):1271–3.
71. Salerno J, Knoppers BM, Lee LM, et al. Ethics, big data and computing in epidemiology and public health. *Ann Epidemiol* 2017 May;**27**(5):297–301.