



Published in final edited form as:

Nat Med. 2019 January ; 25(1): 30–36. doi:10.1038/s41591-018-0307-0.

The practical implementation of artificial intelligence technologies in medicine

Jianxing He^{1,2,9,*}, Sally L. Baxter^{3,4,5,9}, Jie Xu⁶, Jiming Xu⁷, Xingtao Zhou⁸, Kang Zhang^{1,2,3,5,*}

¹Department of Thoracic Surgery/Oncology, First Affiliated Hospital of Guangzhou Medical University, China State Key Laboratory and National Clinical Research Center for Respiratory Disease, Guangzhou, China.

²Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou, China.

³Shiley Eye Institute and Institute of Engineering in Medicine, University of California San Diego, La Jolla, CA, USA.

⁴Department of Biomedical Informatics, University of California San Diego, La Jolla, CA, USA.

⁵San Diego Veterans Affairs Health System, La Jolla, CA, USA.

⁶Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Science Key Lab, Beijing, China.

⁷Yidu Cloud Technology Inc., Beijing, China.

⁸Department of Ophthalmology, Eye and ENT Hospital, Fudan University, Shanghai, China.

⁹These authors contributed equally: Jianxing He, Sally L. Baxter.

Abstract

The development of artificial intelligence (AI)-based technologies in medicine is advancing rapidly, but real-world clinical implementation has not yet become a reality. Here we review some of the key practical issues surrounding the implementation of AI into existing clinical workflows, including data sharing and privacy, transparency of algorithms, data standardization, and interoperability across multiple platforms, and concern for patient safety. We summarize the current regulatory environment in the United States and highlight comparisons with other regions in the world, notably Europe and China.

AI has recently experienced an era of explosive growth across many industries, and the healthcare industry is no exception. Studies across multiple medical specialties have employed AI to mimic the diagnostic abilities of physicians^{1–4}. The hope is that AI may

*Correspondence should be addressed to J.H. or K.Z., hejx@vip.163.com; kang.zhang@gmail.com.

Competing Interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

augment the ability of humans to provide healthcare. However, although these technologies are rapidly advancing, their implementation into patient-care settings has not yet become widespread. Here we review some key issues surrounding implementation of AI-based technologies in healthcare.

The current status of AI in medicine

AI involves the development of computer algorithms to perform tasks typically associated with human intelligence⁵. AI is broadly used in both the technical and popular lexicon to encompass a spectrum of learning, including but not limited to machine learning, representation learning, deep learning, and natural language processing (see Box 1 for definitions). Regardless of the specific technique, the general aim of these technologies in medicine is to use computer algorithms to uncover relevant information from data and to assist clinical decision-making⁶. AI technologies can perform a wide array of functions, such as aiding in diagnosis generation and therapy selection, making risk predictions and stratifying disease, reducing medical errors, and improving productivity^{7,8} (Fig. 1).

Personal health data may include demographics, healthcare provider notes, images, laboratory results, genetic testing data, and recordings from medical devices or wearable sensors. A multitude of technology platforms may be involved in the generation or collection of this data, including network servers, electronic health records (EHRs), genomic data, personal computers, smart phones and mobile apps, and wearable devices and sensors. With improved global connectivity via the Internet and technology with cloud-based capabilities, data access and distribution has become easier as well. Integration of big data concerning health and disease will provide unprecedented opportunities in the management of healthcare information at the interface of patients, physicians, hospitals and healthcare authorities, and regulatory bodies (Fig. 2).

Potential roles for AI in clinical settings.

We envision several ways in which AI-based technologies could be implemented into clinical practice. The first is as a triage or screening tool. For example, AI could analyze radiology images and use the probability of disease as a basis to decide which images should be interpreted first by a human radiologist⁹, or it could examine retinal images to determine which patients have a vision-threatening condition and should be referred to an ophthalmologist². Similarly, the Babylon app, an AI-based chatbot being piloted in the United Kingdom¹⁰, is essentially a triage tool used to differentiate patients who simply need reassurance from those who need referral for an in-person examination. AI-based triaging would theoretically decrease burden on the healthcare system and direct resources toward the patients most likely to have a real medical need.

AI technologies may also function in a replacement scenario. While unlikely to replace human healthcare providers entirely, AI may perform certain tasks with greater consistency, speed, and reproducibility than humans. Examples include estimation of bone age on radiographic exams⁹, diagnosing treatable retinal diseases on optical coherence tomography^{2,11}, or quantifying vessel stenosis and other metrics on cardiac imaging¹². By automating tasks which are not theoretically complex but can be incredibly labor- and time-

intensive, healthcare providers may be freed to tackle more complex tasks, representing an improved use of human capital.

Perhaps the most powerful role for AI will be as an add-on to or augmentation of human providers. Studies have demonstrated a synergistic effect when clinicians and AI work together, producing better results than either alone^{13,14}. AI-based technologies could also augment real-time clinical decision support, resulting in improved efforts toward precision medicine¹⁵.

Engagement in the AI health space.

The AI health space seems to be constantly expanding. The National Science and Technology Council's Committee on Technology estimated that the US government's investment in research and development in AI-related technologies was ~\$1.1 billion in 2015 and that this will substantially increase over subsequent years¹⁶. While this amount also includes funding for other areas, even within healthcare the numbers are staggering. To illustrate, Massachusetts General Hospital and Brigham and Women's Hospital have spent more than \$1 billion on health information and data collection infrastructures to promote operationalizing AI algorithms for real-world clinical practice¹⁷. While much of this investment was for an EHR system, because many AI technologies will be expected to interface with the EHR in order to influence the delivery of healthcare, this infrastructure investment can be considered an initial step toward translating AI algorithms, even if EHR implementation is not AI per se. In addition, academic institutions are increasingly engaged in AI-related research: 212 AI-related publications were indexed globally in 1990, increasing to 1,153 in 2014¹⁸.

On the industry side, AI-related health start-ups have experienced increased funding and investor interest. By 2021, growth in the AI market is expected to reach \$6.6 billion¹⁹. The number of healthcare-focused AI deals increased from less than 20 in 2012 to nearly 70 by mid-2016²⁰. Even companies not traditionally linked with healthcare, such as Google and Amazon, are increasing their involvement in this sector.

Despite this growing interest in healthcare-related AI, substantial translation or implementation of these technologies into clinical use has not yet transpired. How we tackle issues in implementation in the next few years will likely have far-reaching impacts for the future practice of medicine.

Key issues in implementation

Above and beyond building AI algorithms, 'productizing' them for clinical use is incredibly complex¹⁷. This productization process requires the availability of a massive amount of data, integration into complex existing clinical workflows, and compliance with regulatory frameworks.

Data sharing.

Not only are data necessary for initial training, a continued data supply is needed for ongoing training, validation, and improvement of AI algorithms. For widespread

implementation, data may need to be shared across multiple institutions and potentially across nations. The data would need to be anonymized and deidentified, and informed consent processes would need to include the possibility of wide distribution. With this scale of dissemination, the notions of patient confidentiality and privacy may need to be reimagined entirely²¹. Subsequently, cybersecurity measures will be increasingly important for addressing the risks of inappropriate use of datasets, inaccurate or inappropriate disclosures, and limitations in deidentification techniques²².

Existing data-sharing efforts include biobanks and international consortia for medical imaging databases, such as the Cardiac Atlas Project²³, the Visual Concept Extraction Challenge in Radiology (VISCERAL) Project²⁴, the UK Biobank (<http://www.ukbiobank.ac.uk/>), and the Kaggle Data Science Bowl (<https://www.kaggle.com/datasets>). However, the extent of data sharing required for widespread adoption of AI technologies across health systems will require more expansive efforts.

The current healthcare environment holds little incentive for data sharing⁷. This may change with ongoing healthcare reforms that favor bundled-outcome-based reimbursement over fee-for-service models. This would create greater incentive to compile and exchange information⁷. Furthermore, the government should promote data sharing. The National Science and Technology Council Committee on Technology recommended that open data standards for AI should be a key priority for federal agencies¹⁶. Similarly, a strategic plan issued by the Obama Administration envisioned a ‘sociotechnical’ infrastructure with a wide variety of accessible training and testing datasets²², and the current Trump administration has also expressed support for research and development in AI²⁵.

Some have proposed creating anonymized ‘benchmarking data-sets’ with known diagnoses that are updated and ‘calibrated’ at regular intervals using local data from the implementing institutions, similar to how clinical laboratories maintain a local reference standard for blood-based biomarkers⁹. These maintenance measures would clearly require extensive data sharing and significant human effort. The local calibration is important, because certain algorithms may have local or cultural-specific parameters that may not be generalizable to different populations.

Transparency.

Transparency of data and AI algorithms is also a major concern. Transparency is relevant at multiple levels. First, in the case of supervised learning (Box 1), the accuracy of predictions relies heavily on the accuracy of the underlying annotations inputted into the algorithm. Poorly labeled data will yield poor results²⁶, so transparency of labeling such that others can critically evaluate the training process for a supervised learning algorithm is paramount to ensuring accuracy.

Besides issues in annotation or labeling, transparency also relates to model interpretability—in other words, humans should be able to understand or interpret how a given technology reached a certain decision or prediction²⁷ (<https://christophm.github.io/interpretable-ml-book/>). This kind of ‘transparent’ AI contrasts the ‘opaque’ or ‘black box’ approach²⁸. It is thought that if the system’s reasoning can be explained, then humans can verify whether the

reasoning is sound. If the system's reasoning cannot be explained, such evaluation would not be practically feasible. However, there are tradeoffs—in certain instances, enforcing transparency and interpretability can potentially result in decreased accuracy or predictive performance of a model. And certain scenarios may not require interpretability, such as if the model has no significant impact or if the problem is well-studied (<https://christophm.github.io/interpretable-ml-book/>). It is obvious that the delivery of healthcare does not fall into these categories. AI technologies will need transparency to justify a particular diagnosis, treatment recommendation, or outcome prediction²². Furthermore, only with transparency can physicians and researchers perform analyses that could potentially uncover new clinical insights.

Another reason why transparency is important is that AI technologies have the potential for algorithmic bias, thereby reinforcing discriminatory practices based on race, sex, or other features^{21,22}. Transparency of training data and of model interpretability would allow examination for any potential bias. Ideally, machine learning might even help resolve healthcare disparities if designed to compensate for known biases²¹.

Transparency will be difficult to achieve if companies purposefully make their algorithms opaque for proprietary or financial reasons. Physicians and other stakeholders in the healthcare system must demand transparency to facilitate patient safety.

Patient safety.

What mechanisms will (or should) be in place for quality control of algorithms remain unclear. The US Food and Drug Administration (FDA) recognizes that AI-based technologies are distinct from traditional medical devices. Along with the International Medical Device Regulators Forum (IMDRF), the FDA has defined a new category called Software as Medical Device (SaMD) and has expressed the need for an updated regulatory framework, the details of which we will explore in the next section.

Some have proposed that AI technologies should be held to the same standard as clinical laboratories—that is, local standards and established minimum performance metrics for critical abnormalities should be in place⁹. AI systems face safety challenges in the forms of complex environments, periods of learning during which the system's behavior may be unpredictable, and the uncertainty of human-machine interactions, which may result in significant variation in the system's performance.

Another issue related to patient safety is that of accountability. If a patient suffers an adverse event due to an AI-based technology, who is responsible? AI technologies will undoubtedly alter the traditional physician-patient relationship. Inherent in that change is a potential shift in the physician's sense of personal responsibility. If AI indeed completely replaces certain tasks previously performed by the physician, perhaps that shift in responsibility is warranted. But if the physician is not responsible, who is? One could reasonably propose multiple sources—the vendor providing the software platform, the developer who built the algorithm, or the source for the training data. The patient safety movement is already shifting away from blaming individual 'bad actors' and working toward identifying systems-wide issues as opportunities for improvement and reduction in potentially avoidable adverse events. The

same principles could be applied to AI technology implementation, but where liability will ultimately rest remains to be seen.

Data standardization and integration into existing clinical workflows.

Data standardization is critical for implementation. Data standardization refers to the process of transforming data into a common format that can be understood across different tools and methodologies. This is a key concern because data are collected in different methods for different purposes and can be stored in a wide range of formats using variable database and information systems. Hence, the same data (e.g., a particular biomarker such as blood glucose) can be represented in many different ways across these different systems. Healthcare data has been shown to be more heterogeneous and variable than research data produced within other fields²⁹. In order to effectively use these data in AI-based technologies, they need to be standardized into a common format. With the complexity of healthcare data and the massive volumes of patient information, data standardization should occur at the initial development stage and not at the user end.

Interoperability will be essential given the multiple components of a typical clinical workflow. For example, Tang et al.⁹ posited that for an AI-assisted radiology workflow, algorithms for protocolling, study prioritization, feature analysis and extraction, and automated report generation could each conceivably be a product of individual specialized vendors. A set of standards would be necessary to allow integration between these different algorithms and also to allow algorithms to be run on different equipment. Without early efforts to optimize interoperability, the practical effectiveness of AI technologies will be severely limited.

The Digital Imaging and Communications in Medicine (DICOM) standards and the picture archiving and communication system (PACS) revolutionized medical imaging by providing a consistent platform for data management. A similar set of standards should be applied to AI-based technologies to develop a consistent nomenclature to facilitate consistent methods of data storage and retrieval.

One such set of standards that is rapidly evolving worldwide for the purpose of clinical translation is the Fast Healthcare Interoperability Resources (FHIR) framework (<https://www.hl7.org/fhir/summary.html>). This framework is based on Health Level 7 (HL7), a framework of standards for electronic health information exchange, and is geared toward implementation. FHIR utilizes a set of modular components, known as ‘Resources,’ which can be assembled into working systems that will facilitate data sharing within the EHR and mobile-based apps as well as cloud-based communications. Looking to the future, we anticipate that the FHIR framework will be critical for implementation of AI-based technologies in the healthcare sector that utilize electronic data, just as DICOM and PACS have become critical for exchange of digital medical images.

Financial issues in AI technology implementation.

As mentioned previously, optimal performance of AI systems will require ongoing maintenance not only in incorporation of increasing amounts of patient data, but also in terms of updating software algorithms and ensuring hardware operability. Equipment

upgrades may also be needed to support software updates. All of this maintenance activity will require not only significant effort in human capital, but also a funding support mechanism. Funding will be critical to ensuring successful implementation and ongoing process improvement, and currently it is not clear how use of AI technologies will be reimbursed.

On the other hand, there is also potential for abuse by AI technology developers. For example, clinical decision support systems could be programmed to increase profits for certain drugs, tests, or devices without clinical users being aware of this manipulation²¹. For all medical devices, a tension exists between providing ethical medical care and generating profit²¹. AI technologies will not be immune to that tension, and it should be openly acknowledged and addressed during implementation processes.

Another concern that relates to the financial aspect of implementation is whether there is sufficient business incentive to motivate translation of these technologies. While business incentives are by no means the only way to advance healthcare, historically they have played a key role in facilitating change. The question of whether AI-based technologies actually bring added value to healthcare with improved outcomes is unanswered. The notion of advanced technology with incredible potential that is not yet fully realized is not new. For example, gene therapy, genomic-driven personalized medicine, and EHRs are all technologies that were purported to deliver revolutionary improvements in the delivery of healthcare, but thus far many have felt that their potential has exceeded their performance. However, those fields are delivering ongoing advancements with continuing promise for the future. Similarly, the application of AI-based technologies to medicine is still in its early stages. While the initial investments from government, academia, and industry are growing, whether these will sustain into the future remains to be seen and in part may depend on the successes of early translational efforts.

Education of an AI-literate workforce.

All stakeholders need to be actively involved in the AI implementation process in medicine. The workforce should be educated about AI and understand both its benefits and limitations. Medical school and residency curricula should be updated to include topics in health informatics, computer science, and statistics⁹. In an ideal world, physicians would understand the construction of algorithms, comprehend the datasets underlying their outputs, and, importantly, understand their limitations. But in a world of finite resources and competing demands on clinicians' time, it may not be reasonable to expect every provider to reach that level of understanding. However, as AI becomes increasingly prevalent not only in healthcare, but also in our world at large, we anticipate that an increasing number of clinicians will be interested in obtaining this type of dedicated training, whether through research projects during medical school, residency, or fellowship, or through dedicated training programs in informatics, such as those offered by the National Library of Medicine. Ultimately healthcare providers will need this knowledge to maximize their functioning on human-machine teams. Additionally, as patient advocates, a cadre of healthcare workers needs to understand these technologies in order to educate policymakers on the complexities of clinical decision-making and the consequences of potential misuse.

AI policy and the regulatory environment in the united States

One key element to implementing AI technologies in healthcare is the development of regulatory standards for assessment of safety and efficacy. Conventional medical device approvals in the United States currently go through the FDA's Center for Devices and Radiological Health (CDRH) and are classified as Class I, II, or III, with higher numbers indicating increasing risk of illness or injury and therefore a higher level of regulatory control to ensure safety and effectiveness (<https://www.fda.gov/ForPatients/Approvals/Devices/ucm405381.htm>). Medical devices undergo an extensive regulatory process, which includes establishment registration, medical device listing, and a Premarket Notification 510(k) for Class I or II devices or Premarket Approval (PMA) for Class III devices. The process is very complicated and can be lengthy. The average time for FDA 510(k) submissions to be cleared in 2015 was 177 days³⁰. PMA reviews are normally longer than 180 days (<https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketnotifications/premarket-approvalpma/>). Evaluations are performed for individual medical devices, and any changes to the devices after approvals require an amendment or a supplement, resulting in another review cycle of 180 days (<https://www.fda.gov/MedicalDevices/ucm050467.htm>). The process is not well-suited for fast-paced cycles of iterative modification, which is what software entails.

The International Medical Device Regulators Forum (IMDRF) has defined AI-based technologies as being encompassed by the term Software as a Medical Device (SaMD)³¹, and the FDA has put forth the Digital Health Innovation Action Plan³², which outlines its approach to SaMD, including AI technologies. A few AI technologies have already received FDA approval (Box 2). The FDA has also developed a Software Precertification Program³³ (Fig. 3), a 'more streamlined and efficient' voluntary pathway resulting in improved access to technologies. The FDA acknowledges that traditional forms of medical-device regulation are not well suited for the faster pace of design and modification for software-based medical devices.

According to the FDA's description of the program, this pathway is organization-centric and focuses on the technology developer instead of the individual products. Manufacturers who demonstrate 'organizational excellence' can achieve precertification, facilitating exemption from premarket review for lower-risk SaMD products or faster review for higher-risk SaMD products. These organizations are expected to continue monitoring the safety and effectiveness of their products with real-world performance data, some of which may be automatable. The FDA predicts that with this strategy, precertification and postmarket oversight will allow low-risk devices to forego premarket review and be introduced directly to the market. In addition, the overall timeline for a precertified company's SaMD product review should be shorter than traditional processes. The FDA's hope is that this approach will support continued innovation, allow increased availability of new and updated software, and better focus the organization's resources on higher-risk developers and products.

In our view, the FDA's efforts to define a completely new process for this type of technology acknowledges the delicate and difficult balance between protecting patients and promoting innovation. With overregulation, innovation and advancements in this field may be stifled. However, with under-regulation, there may be risks to patient safety.

The FDA invited public comment on the program throughout the spring of 2018, which led to some additional clarifications³³. The FDA will continue to actively review public feedback throughout this iterative process. We expect that these policies will continue to evolve as these technologies undergo real-world implementation and that this will be an area of dynamic change in the coming years.

The first ‘autonomous’ AI diagnostic system to receive FDA approval for marketing was IDx-DR (Box 3). Implementation of the device has begun at the University of Iowa for screening patients with diabetes³⁴. This represents one of the earliest translations of an autonomous AI screening system into real-world clinical practice, and it will be interesting to see what kind of impact the screening system has and whether it actually leads to improvement in visual outcomes for patients with diabetes.

The view toward global implementation of AI in healthcare

The issues surrounding AI implementation are universal in nature and are not unique to the United States. One of the organizations leading guidance in this area is the International Medical Device Regulators Forum (IMDRF), a voluntary group of medical device regulators working toward harmonizing international medical device regulation (<http://www.imdrf.org/>). Its members currently include Australia, Brazil, Canada, China, Europe, Japan, Russia, Singapore, South Korea, and the United States. The IMDRF has issued key definitions for SaMD and a common framework for regulators³¹. Its guidance also recommends a continuous iterative process based on real-world performance data³⁵. Similar to the FDA approach, the IMDRF states that low-risk SaMD may not require independent review.

The European Union (EU) has issued some key regulatory policies that will affect AI implementation in their healthcare systems. The General Data Protection and Regulation (GDPR) outlines a comprehensive set of regulations for the collection, storage, and use of personal information, which was adopted by the European Parliament in April 2016 and became effective in May 2018³⁶. A critical component of this regulation is Article 22, “Automated individual decision making, including profiling.” This article of regulation describes the right of citizens to receive an explanation for algorithmic decisions. This would potentially prohibit an array of algorithms currently in use in advertising and social networks³⁶. Although the GDPR applies to data from EU citizens, given the global nature of the Internet and modern connectivity, the regulation has impacted companies from the US and worldwide as they adapt to the new regulations.

The GDPR will affect AI implementation in healthcare in several ways. First, it requires explicit and informed consent before any collection of personal data. Informed consent has been a long-standing component of medical practice (unlike in social media or online-based marketing), but having to obtain informed consent for any collection of data still represents a higher bar than obtaining consent for specific items, such as procedures or surgical interventions. Second, the new regulation essentially lends power to the person providing the data to track what data is being collected and to be able to request removal of their data. In the healthcare context, this will shift some of the power balance toward the patient and highlights the importance of ongoing work needed to protect patient privacy and to

determine appropriate governance regarding data ownership. Finally, the need for a ‘right to explanation’ will potentially limit the types of models that manufacturers are able to use in health-related applications. However, as we described earlier in this review, model interpretability is important in AI-based healthcare applications given the high stakes of dealing with human health, and this requirement may actually help AI applications become more reliable and trustable. In addition, this requirement of a right to explanation will hold the manufacturers of AI-based technologies more accountable.

The GDPR regulation has occurred so recently that the long-term effects remain to be seen. Companies may need to enact significant innovations in their practices in order to meet the new standards. The GDPR regulation may also eventually affect regulatory practices in other countries. We anticipate that the GDPR may potentially slow down AI implementation in healthcare in the short term owing to the stricter regulatory standards, but it may facilitate implementation over the long term by promoting public trust and patient engagement.

China is also a major player in AI in the global arena. With the world’s largest population and a relatively centralized healthcare system, data for training and validation of AI algorithms are vast. Healthcare equity has been a major concern due to the uneven distribution of resources between urban and rural areas. As such, AI technologies represent an opportunity to overhaul China’s medical system. The rapid adoption of mobile technology and increased Internet connectivity throughout China is projected to facilitate adoption of AI technologies, which could help make more efficient triage and referral systems for patients. Furthermore, Chinese health providers envision that the combination of AI technologies with wearable devices could assist with health maintenance as well as disease surveillance on a broad scale.

In July 2017, the China State Council issued a development plan for AI, including strong promotion of AI applications in the healthcare system. This plan issued favorable policies and funding for AI start-up companies and implementation efforts. Hundreds of new start-up companies working on AI applications in healthcare in China have emerged. Similar to how Google and Amazon are expanding into the health space in the United States, in China, Internet giants such as Alibaba, Tencent, and Baidu have also established their own groups to work on healthcare AI applications.

In terms of implementing AI technologies in real clinical practice, AI-based screening tools have already been deployed in clinical trials in multiple Chinese hospitals. These include AI-based tools for the diagnosis of lung cancer, esophageal cancer, diabetic retinopathy, and general diagnostic assistance in pathology examinations. The first AI-powered medical device is expected to be approved by the Chinese Food and Drug Administration in the first half of 2019.

One example of a successful implementation in China is an AI-based screening and referral system for the diagnosis and referral of major eye and systemic diseases in the Kashi First People’s Hospital and its healthcare network. Located in the remote most western part of the Xinjiang autonomous region of China, this health system serves a population of 4.5 million people scattered across a mountainous area of 112,057 km². This population has a high

prevalence of blinding eye diseases, including diabetic retinopathy, glaucoma, retinal vein occlusions, and age-related macular degeneration, as well as common chronic systemic diseases including diabetes, hypertension, cardiovascular disease, and stroke. This AI system employs retinal photographs captured by nonmydriatic fundus cameras to screen for and diagnose the aforementioned diseases. Preliminary results demonstrate high accuracy of the AI-generated diagnoses, comparable to that of a trained eye doctor.

Future directions

We anticipate that the fields that will see the earliest translation of AI-based technologies are those with a strong image-based or visual component that is amenable to automated analysis or diagnostic prediction—these include radiology, pathology, ophthalmology, and dermatology. Ophthalmology has seen the first FDA clearance for an autonomous screening tool, and more SaMDs are anticipated to come through the pipeline in the near future. Although fields requiring integration of many different types of data, such as internal medicine, or fields that have a prominent procedural component, such as the surgical specialties, may require a longer time before fully realizing the operationalization of AI-based technology, research in AI-related applications across all specialties of medicine is rapidly advancing.

The implementation of AI-based technologies in healthcare will provide no shortage of work for the future. Specialties may decide to create organizations specifically oriented toward AI implementation, as the American College of Radiology has done with the creation of its Data Science Institute³⁷. Having specific task force committees to deal with AI implementation issues may be useful for developing a common vision at a specialty-wide level.

Ongoing research will be needed to develop new AI algorithms for medical applications and to improve upon existing ones. In particular, interdisciplinary collaborations will be crucial to ensure that the goals of programmers developing the algorithms correspond with the goals of clinicians providing patient care. Development of a workforce that is cross-trained such that communication and collaboration are possible between physicians and healthcare providers, data scientists, computer scientists, and engineers will be crucial. While these technologies hold promise for increasing productivity and improving outcomes, it must be remembered that they, just like their human creators, are not infallible. It is necessary to evaluate and implement them with a critical eye, keeping in mind their limitations and educating policymakers to do the same.

Acknowledgements

This study was funded in part by the Innovative team (B185004102) and Backbone talent (B185004075) training program for high-level universities of Guangzhou Medical University, Guangzhou Regenerative Medicine and Health Guangdong Laboratory.

References

1. Gulshan V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316, 2402 (2016). [PubMed: 27898976]

2. Kermany DS et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172, 1122–1131.e9 (2018). [PubMed: 29474911]
3. Esteva A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115–118 (2017). [PubMed: 28117445]
4. Cheng J-Z et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci. Rep.* 6, 24454 (2016).
5. Goodfellow I, Bengio Y & Courville A. *Deep Learning* (The MIT Press, Cambridge, MA, USA, 2016).
6. Murdoch TB & Detsky AS The inevitable application of big data to health care. *JAMA* 309, 1351–1352 (2013). [PubMed: 23549579]
7. Jiang F. et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc. Neurol.* 21, 230–243 (2017).
8. Johnson KW et al. Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.* 71, 2668–2679 (2018). [PubMed: 29880128]
9. Tang A. et al. Canadian association of radiologists white paper on artificial intelligence in radiology. *Can. Assoc. Radiol. J. J. Assoc. Can. Radiol.* 69, 120–135 (2018).
10. Pelcyger B. Artificial intelligence in healthcare: Babylon Health & IBM Watson take the lead. *Prescouter* <https://prescouter.com/2017/12/artificial-intelligence-healthcare/> (2017).
11. De Fauw J. et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* 24, 1342–1350 (2018). [PubMed: 30104768]
12. Slomka PJ et al. Cardiac imaging: working towards fully-automated machine analysis & interpretation. *Expert Rev. Med. Devices* 14, 197–212 (2017). [PubMed: 28277804]
13. Wang D, Khosla A, Gargeya R, Irshad H & Beck AH Deep learning for identifying metastatic breast cancer. Preprint at <https://arxiv.org/abs/1606.05718> (2016).
14. Lakhani P & Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* 284, 574–582 (2017). [PubMed: 28436741]
15. Sitapati A. et al. Integrated precision medicine: the role of electronic health records in delivering personalized treatment. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 9, 1–12 (2017).
16. Executive Office of the President & National Science and Technology Council Committee on Technology. *Preparing for the Future of Artificial Intelligence* (2016).
17. Dreyer KJ & Geis JR When machines think: radiology's next frontier. *Radiology* 285, 713–718 (2017). [PubMed: 29155639]
18. Niu J, Tang W, Xu F, Zhou X & Song Y. Global research on artificial intelligence from 1990–2014: spatially-explicit bibliometric analysis. *ISPRS Int. J. Geo-Inf.* 5, 66 (2016).
19. Collier M, Fu R & Yin L. *Artificial Intelligence: Healthcare's New Nervous System* (Accenture, 2017).
20. From Virtual Nurses To Drug Discovery: 106 Artificial Intelligence Startups InHealthcare. CB Insights Research <https://www.cbinsights.com/research/artificial-intelligence-startups-healthcare/> (2017).
21. Char DS, Shah NH & Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N. Engl. J. Med.* 378, 981–983 (2018). [PubMed: 29539284]
22. National Science and Technology Council Committee on Technology Council & Networking and Information Technology Research and Development Subcommittee. *The National Artificial Intelligence Research and Development Strategic Plan* (2016).
23. Fonseca CG et al. The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart. *Bioinforma. Oxf. Engl.* 27, 2288–2295 (2011).
24. Jimenez-Del-Toro O. et al. Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms: VISCERAL anatomy benchmarks. *IEEE Trans. Med. Imaging* 35, 2459–2475 (2016). [PubMed: 27305669]
25. Artificial Intelligence for the American People. The White House <https://www.whitehouse.gov/briefings-statements/artificial-intelligence-american-people/> (2018).

26. Hashimoto DA, Rosman G, Rus D & Meireles OR Artificial intelligence in surgery: promises and perils. *Ann. Surg.* 268, 70–76 (2018). [PubMed: 29389679]
27. Patrzyk PM, Link D & Marewski JN Human-like machines: transparency and comprehensibility. *Behav. Brain Sci.* 40, e276 (2017). [PubMed: 29342707]
28. Sussillo D & Barak O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput.* 25, 626–649 (2013). [PubMed: 23272922]
29. Kruse CS, Goswamy R, Raval Y & Marawi S. Challenges and opportunities of big data in health care: a systematic review. *JMIR Med. Inform.* 4, e38 (2016). [PubMed: 27872036]
30. How long it takes the US FDA to approve 510(k) submissions. *Emergo* <https://www.emergobyul.com/resources/research/fda-510k-review-times-research> (2013).
31. IMDRF SaMD Working Group. Software as a Medical Device (SaMD): Key Definitions (International Medical Device Regulators Form, 2013).
32. US Food & Drug Administration & Center for Devices & Radiological Health Digital Health Program. Digital Health Innovation Action Plan (United States Food and Drug Administration, 2018).
33. U.S. Food & Drug Administration. Developing Software Precertification Program: A Working Model (United States Food & Drug Administration, 2018).
34. IDx. University of Iowa Health Care First to Adopt IDx-DR in a Diabetes Care Setting. *Cision PR Newswire* <https://www.prnewswire.com/news-releases/university-of-iowa-health-care-first-to-adopt-idx-dr-in-a-diabetes-care-setting-300672070.html> (2018).
35. Software as a Medical Device Working Group. Software as a Medical Device (SaMD): Clinical Evaluation (United States Food and Drug Administration, 2017).
36. Goodman B & Flaxman S. European Union regulations on algorithmic decision-making and a ‘right to explanation. *AI Mag.* 38, 50 (2017).
37. Thakar S. 4 ways the ACR’s Data Science Institute is looking to implement AI in clinical practice. *Radiology Business* <http://www.radiologybusiness.com/topics/artificial-intelligence/4-ways-acrs-data-science-institute-looking-implement-ai-clinica> (2018).

Box 1 |**Key terms in artificial intelligence**

- Artificial intelligence: A branch of applied computer science wherein computer algorithms are trained to perform tasks typically associated with human intelligence.
- Machine learning: Providing knowledge to computers through data without explicit programming. Attempts to optimize a ‘mapping’ between inputs and outputs.
- Representation learning: Learning effective representations of a data source algorithmically, as opposed to hand-crafting combinations of data features.
- Deep learning: Multiple processing layers are used to learn representations of data with multiple layers of abstraction.
- Supervised learning: Training is conducted with specific labels or annotations.
- Unsupervised learning: Training is conducted without any specific labels, and the algorithm clusters data to reveal underlying patterns.
- Natural language processing: The organization of unstructured narrative text into a structured form that can be interpreted by a machine and allows for automated information extraction.

Box 2 |**Examples of artificial intelligence technologies that have received FDA approval in the United States**

- **Arterys:** Aids in finding lesions within pulmonary computed tomography (CT) scans and liver CT and magnetic resonance imaging (MRI) scans using AI to segment lesions and nodules. This is the first FDA-approved deep learning clinical platform.
- **IDx-DR:** Provides automatic detection of more than mild diabetic retinopathy in adults 22 years of age or older diagnosed with diabetes who have not been previously diagnosed with diabetic retinopathy. This is meant to be used in primary care settings with subsequent referral to an eye specialist if indicated and is the first autonomous AI diagnostic system, as no clinician interpretation is needed before a screening result is generated.
- **Guardian Connect (Medtronic):** Continuously monitors glucose and sends collected data to a smartphone app. This uses International Business Machines (IBM) Watson technology to predict major fluctuations in blood glucose levels 10–60 minutes in advance.

Box 3 |**The case study of IDx-DR**

The case of IDx-DR highlights one of the earliest successes of an AI-based technology completing the regulatory process in the United States. IDx-DR is a software program designed to perform screening for diabetic retinopathy at primary care offices, or more generally, in contexts where a provider specifically trained in eye care (e.g., an optometrist or ophthalmologist) is not readily available. The software program analyzes digital color photographs of a patient's retinas using an AI algorithm to provide a screening result—either (i) “more than mild diabetic retinopathy detected; refer to an eye care professional” or (ii) “negative for more than mild diabetic retinopathy; rescreen in 12 months” (<https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm596575.htm>). Although other devices with an AI component have been previously approved (Box 2), this is the first to provide a screening result without an image interpretation provided by a clinician: it is the first autonomous AI diagnostic system. Furthermore, instead of going through the traditional pathway, IDx-DR underwent the Automatic Class III or De Novo premarket pathway and also achieved Breakthrough Device designation. The De Novo pathway is for novel medical devices for which general controls provide reasonable assurance of safety and effectiveness (<https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYour-Device/PremarketSubmissions/ucm462775.htm>); essentially, it is an alternative regulatory pathway for low- to moderate-risk devices. In the review process, the FDA based its clearance on the performance of the algorithm in a clinical trial of 900 patients (NCT02963441), which was conducted at ten primary care sites across the United States. The FDA was closely involved with advising and guiding the company throughout the clinical trial and in defining meaningful endpoints. This early involvement and active collaboration between the device manufacturer and the FDA undoubtedly facilitated the review. Early engagement with the FDA and pursuance of the De Novo pathway likely represents an advantageous strategy for other manufacturers of AI-based technologies to consider as well.

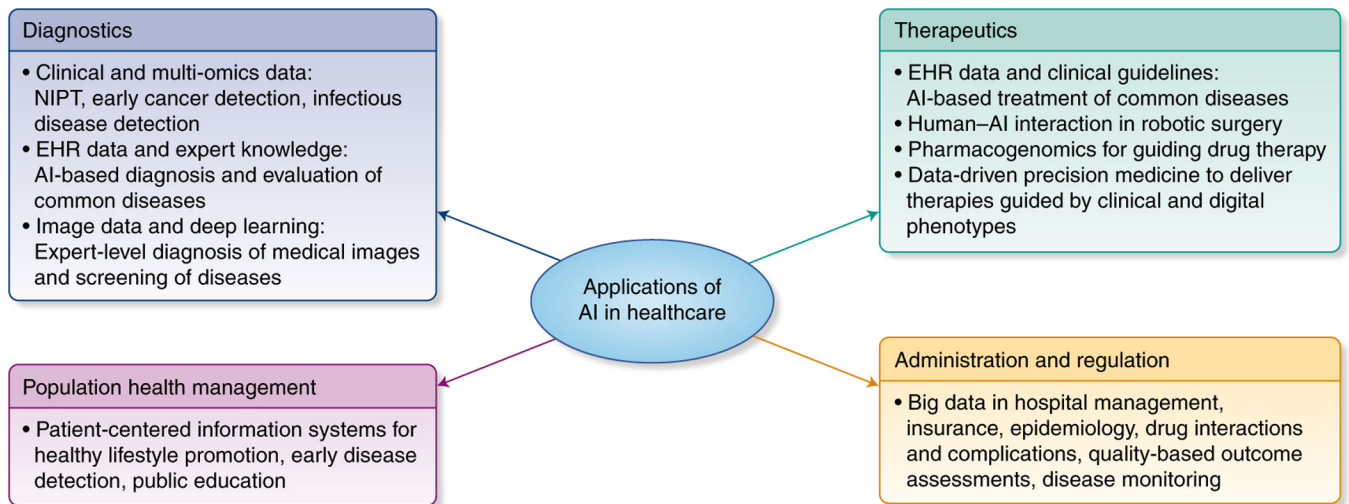


Fig. 1 |. Potential roles of AI-based technologies in healthcare.

In the healthcare space, AI is poised to play major roles across a spectrum of application domains, including diagnostics, therapeutics, population health management, administration, and regulation. NIPT, noninvasive prenatal test. Credit: Debbie Maizels/Springer Nature

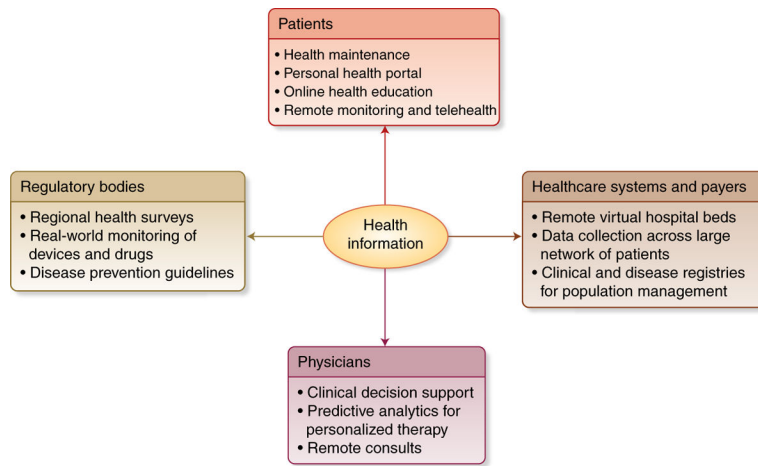


Fig. 2 |. Integration of patient health information at multiple interfaces.

The vast quantity and accessibility of electronic health information will influence decision-making on multiple fronts, including for patients, physicians, healthcare systems, healthcare providers, and regulatory bodies. Standardization of information storage and retrieval will be critical for facilitating information exchange across these multiple interfaces. Credit: Debbie Maizels/Springer Nature

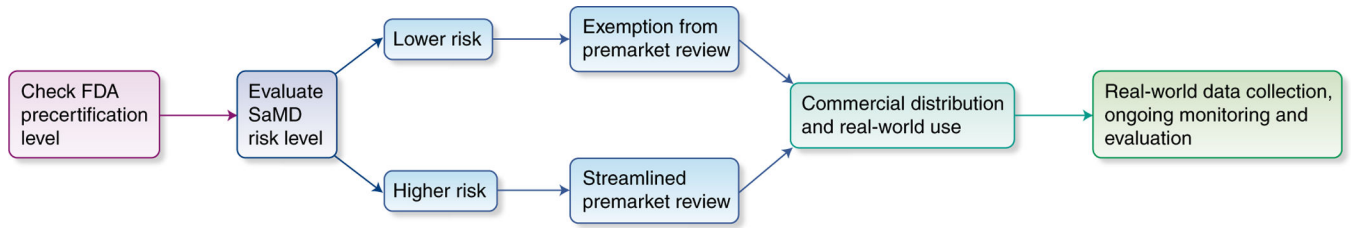


Fig. 3 |. Conceptual diagram of the FDA precertification for SaMD.

This approach represents an organization-centric evaluation to facilitate streamlined review and faster adoption of technology. Credit: Debbie Maizels/Springer Nature