



# Delivering personalized medicine in retinal care: from artificial intelligence algorithms to clinical application

J. Jill Hopkins<sup>a</sup>, Pearse A. Keane<sup>b,c</sup>, and Konstantinos Balaskas<sup>b</sup>

## Purpose of review

To review the current status of artificial intelligence systems in ophthalmology and highlight the steps required for clinical translation of artificial intelligence into personalized health care (PHC) in retinal disease.

## Recent findings

Artificial intelligence systems for ophthalmological application have made rapid advances, but are yet to attain a state of technical maturity that allows their adoption into real-world settings. There remains an 'artificial intelligence chasm' in the spheres of validation, regulation, safe implementation, and demonstration of clinical impact that needs to be bridged before the full potential of artificial intelligence to deliver PHC can be realized.

## Summary

Ophthalmology is currently in a stage between the demonstration of the potential of artificial intelligence and widespread deployment. Next stages include aggregating and curating datasets, training and validating artificial intelligence systems, establishing the regulatory framework, implementation and adoption with ongoing evaluation and model adjustment, and finally, meaningful human–artificial intelligence interaction with clinically validated tools that have demonstrated measurable impact on patient and healthcare system outcomes. Ophthalmologists should leverage the ability of artificial intelligence systems to glean insights from large volumes of multivariate data, and to interpret artificial intelligence recommendations in a clinical context. In doing so, the field will be well positioned to lead the transformation of health care in a personalized direction.

## Video abstract

<http://links.lww.com/COOP/A35>.

## Keywords

artificial intelligence, deep learning, personalized health care

## INTRODUCTION

In recent years, the potential of artificial intelligence – in particular, a technique called deep learning – to transform health care has become increasingly apparent. Ophthalmology has been at the forefront of this revolution and may soon serve as an exemplar for other medical specialties. Catalyzed by upheavals of the global coronavirus pandemic, it seems likely artificial intelligence tools/platforms will allow world-class ophthalmic expertise to be brought increasingly into the community, and even the home. This should allow greatly enhanced outcomes/efficiency for patients and healthcare systems, with improved screening, triage, diagnosis, and monitoring of eye disease [1–3]. Predictive algorithm usage may also be a big step

toward the goal of personalized health care (PHC) for patients by harnessing the power of data to deliver individualized care to patients with retinal

<sup>a</sup>Ophthalmology Personalized Healthcare, Genentech, Inc., South San Francisco, California, USA, <sup>b</sup>Moorfields Eye Hospital NHS Foundation Trust and <sup>c</sup>Institute of Ophthalmology, University College London, London, UK

Correspondence to J. Jill Hopkins, MD, FRCS, Genentech, Inc., 1 DNA Way, MS407b, South San Francisco, CA 94080, USA.

Tel: +1 650 302 9527; e-mail: [hopkins.jill@gene.com](mailto:hopkins.jill@gene.com)

**Curr Opin Ophthalmol** 2020, 31:329–336

DOI:10.1097/ICU.0000000000000677

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

## KEY POINTS

- Implementation of artificial intelligence systems in retinal care provides the opportunity to bring world-class expertise into the community and into the home for the benefit of patients.
- There is an urgent need to bridge the artificial intelligence chasm between proof-of-concept development and real-life patient benefit with high-quality clinical validation and implementation science.
- To fulfill the potential of PHC in ophthalmology, there is a tremendous amount of work to be done across all stages of artificial intelligence solutions development and unprecedented collaboration will be required across health care and technology sectors to achieve this.

disease. Perhaps most importantly, artificial intelligence-assisted scientific discovery may provide novel insights into disease progression and underlying pathophysiology [4<sup>11</sup>], and improve clinical trial efficiency and success rates. Thus, the potential exists for artificial intelligence to meaningfully transform the care of patients with retinal disease at every stage of the patient journey to deliver truly personalized medicine and significantly impact the healthcare ecosystem of the future.

Although recent advances in artificial intelligence-enabled health care in retinal disease have been rapid, with some demonstrating clinical and real-world applicability [5<sup>12</sup>,6], several hurdles remain for real patient benefit to be realized. Despite the hype, artificial intelligence in ophthalmology is still at a nascent stage. It is not yet widely used and has not yet led to patient benefits at scale. There remains a huge ‘artificial intelligence chasm’ that needs to be crossed before scalable deployment, meaningful real-world application, and human–artificial intelligence interaction can be achieved [4<sup>11</sup>,7,8].

## NEED FOR BETTER DATA INFRASTRUCTURE

Before validation and scalable deployment, large high-quality clinical datasets are critical for the training and evaluation of deep-learning models. There is a need to go beyond traditional biostatistical approaches using relatively small amounts of data (e.g., from clinical trials and small registries) to harnessing advanced analytic methods on large real-world datasets. A fundamental requirement is that these training datasets must be representative of those that will be encountered in future clinical

applications [3,9]. They also need to be in a computationally tractable form, appropriately annotated, and in most circumstances, robustly deidentified. These requirements can be challenging and time consuming, and impact utilization of currently available datasets. While data from clinical trials (e.g., HARBOR [10]) and cohort studies (e.g., UK Biobank [11]) tend to be of high quality and contain deep individual patient-level information, these data are not always representative of clinical use cases and may not be of sufficient size or generalizability to achieve state-of-the-art results. Consequently, we have seen increased use of wider datasets derived from routine clinical care, such as the American Academy of Ophthalmology IRIS Registry (Intelligent Research In Sight) [12,13] of electronic health records (EHRs), and ophthalmic imaging bioresources such as the INSIGHT Health Data Research Hub for Eye Health [14] in the United Kingdom. In addition to training, it will be important that large datasets are also available to allow independent benchmarking of artificial intelligence systems and facilitate approval of commercial artificial intelligence systems by regulators. Significant recent progress has also been made in deriving ethical and legal frameworks for such work, both in the United States and elsewhere, with patient and public involvement and transparency about usage being paramount considerations [15–17].

With advances in technology, sources and modalities of data that could be harnessed will also expand [e.g., development of home-based vision monitoring and optical coherence tomography (OCT) [18] systems may allow much more frequent functional and structural retinal disease assessments]. With increasing 5G technology usage, along with cloud computing infrastructure [4<sup>11</sup>,19], personal sensor and hyperlocal environmental data may also be captured and used as the substrate for artificial intelligence models. In coming years, we will also see rapid advances in collection of ‘-omic’ data, from genomics to metabolomics to oculomics. Systems and infrastructure to integrate and manage these data will also be critically important. In the future, ophthalmologists will likely use a range of narrowly focused artificial intelligence tools to integrate this information and to make individualized and informed decisions for their patients [19].

## FROM IDEAS TO ALGORITHMS: EARLY-STAGE DEVELOPMENT OF NOVEL OPHTHALMIC ARTIFICIAL INTELLIGENCE SYSTEMS

Deep learning uses artificial neural networks, named because of their superficial resemblance to

biological neural networks, to discover intricate structure in large datasets. Convolutional neural network usage has proven particularly effective for image classification tasks, and thus much of the early focus of artificial intelligence in ophthalmology has been on classification of retinal photographs/OCT images. These choices have also been driven, in large part, by availability of existing large datasets of these images. As the artificial intelligence field continues to evolve and datasets in other areas become available, it is likely that the range of clinical applications (often termed ‘use cases’) will greatly expand. In the coming years, we are likely to see clinical applications for novel areas, such as generative adversarial networks [20,21], which allow synthetic data generation, and reinforcement learning [22,23], the technology underlying game-playing systems such as AlphaGo. In addition, automated deep-learning platform usage will allow clinical researchers without specialized expertise to scope feasibility of deep learning in healthcare applications [24<sup>11</sup>]. This will likely be a major stimulus to novel use of such systems and the next phase of artificial intelligence ‘industrialization’.

As the translational pipeline for clinical artificial intelligence systems evolves and matures, a crucial aspect will be identification – and robust interrogation – of potential use cases. Identification and optimization of the correct metrics to measure performance/utility of an algorithm for specific-use cases will be needed.

The multiple stakeholders involved in the translational pipeline have potentially different needs and interests. For patients, it will be important to know how use will result in better treatment and outcomes. For clinicians, it could mean how it addresses an unmet need, benefits patients, and affects patient management. For regulators [25,26], it will be necessary to receive clear definitions of intended use with justifiable assumptions made from existing data, and how the system may evolve over time. For payers, it will be important to clarify who the product is intended for use by, in what population, in which disease subtype, and under what circumstances. On the contrary, much initial work in this space has not yet reached this level of consideration.

## TOWARD ROBUST CLINICAL VALIDATION

The first step toward bridging the ‘artificial intelligence chasm’ is robust validation of the algorithm’s performance. If an artificial intelligence system can be used to identify congenital cataract, papilledema, or diabetic retinopathy [27–29], how will such

images be obtained and does that already solve the clinical problem? Closely related to this, for artificial intelligence systems to be implemented, it is essential to establish benchmarks for human diagnostic performance at a range of expertise levels for screening, triage, diagnosis, and future prediction of retinal disease. At the same time, it is also essential to ensure that when artificial intelligence systems are used, they result in improvements in or are on par with human diagnostic performance, rather than decreasing accuracy [30], and that they are adequately generalizable to the clinical population at hand.

Several retrospective studies [2] with large numbers of patients have been used to train and test deep-learning algorithms; however, very few prospective studies have evaluated performance of artificial intelligence systems [31<sup>11</sup>,32,33] in ophthalmology. A hierarchy of validation studies with ascending order of reliability would include retrospective *in-silico* validation using a hospital-based dataset with high prevalence of disease; out-of-sample validation by assessing performance on datasets from different patient populations than the one the algorithm was trained on; prospective observational validation by exposure of the algorithm to data representative of the diversity and variation in real-life clinical practice; and in prospective randomized controlled trials (RCTs), where possible. Retrospective studies performed *in silico* are often based on datasets from hospital settings with high prevalence of disease and are prone to selection bias. Very few algorithms have been externally validated with evaluation of human and artificial intelligence performance on the same dataset [34]. Prospective studies are needed to understand diagnostic accuracy of artificial intelligence systems in real-world settings. In a pivotal prospective trial of an artificial intelligence diagnostic system for diabetic retinopathy, diagnostic accuracy in predicting diabetic retinopathy severity was less than what was observed in retrospective studies, highlighting the need to validate performance of artificial intelligence systems in prospective validation studies with data representative of real-life diversity [31<sup>11</sup>]. However, such studies will not address the issue of clinical effectiveness. Interventional RCTs would potentially be needed to address this key question. To date, there are very few RCTs of artificial intelligence systems [35], and such studies, particularly those aiming to use clinical outcomes and quality-of-life improvements as trial endpoints, may be needed for regulatory approval. EHRs and other observational databases also offer promise in validating deep-learning algorithms by extracting/incorporating real-world clinical data.

Technical maturity of artificial intelligence systems will depend upon performance validation in the general population, with low disease prevalence, versus a patient population; evaluation of positive predictive value of artificial intelligence systems to ensure a surge of false-positive results is avoided; and performance versus human experts, including in complex cases. To this end, frameworks and metrics for artificial intelligence performance reporting need to be established. Ability to fine-tune artificial intelligence models for different use cases and devices (interoperability across different imaging modalities and equipment); rapidly validate it for different patient populations than the ones it was originally trained on; technical integration of artificial intelligence tools and expertise at different levels of the care pathway with different levels of technical infrastructure maturity, including community (high-street optician); and secondary care will also be key to technical maturity of artificial intelligence systems.

Regulatory frameworks are fundamental to achieving safe and effective deployment of artificial intelligence algorithms, and efforts are evolving to ensure appropriate regulatory guidelines are in place as artificial intelligence systems become part of routine care delivery [25,36–38]. However, the current pace of innovation in artificial intelligence and unique nature of deep-learning models [8,35] pose a unique set of challenges; artificial intelligence systems are often designed to be autodidactic, and may also be constantly improved/upgraded by the provider throughout their lifecycles. Guidelines in development for ongoing performance monitoring that factor in anticipated modifications to artificial intelligence systems [39] will support identification of performance deficits over time, and allow continuous learning in real-world settings. The evolving artificial intelligence regulatory framework will have implications for clinical trial planning, ability to generate data outside of clinical trials from real-world evidence in the form of EHRs, and commercial development. Regulatory framework [40] for dealing with relevant ethics issues such as patient privacy and anonymity while handling protected health information, data security, and risk mitigation of artificial intelligence-based clinical prediction and decision support will also be needed. Regulatory requirements (prospective versus retrospective validation) may also differ based on the particular use case. It is imperative that all key stakeholders involved in the development and implementation of artificial intelligence systems collaborate with each other, and work closely with regulators in the regulatory framework development [36–38].

Regulation around conventional therapies typically requires postmarketing surveillance and clinical follow-up plans [41,42]. Currently, there are no long-term follow-up data available, with artificial intelligence systems being relatively novel. Artificial intelligence systems may also be prone to effects of the dataset shift [35] brought on by changes in practice or new data sources such as long-term follow-up data. Postmarket monitoring/follow-up and reporting will help implementation of corrective plans such as quantification of performance over time, updating, and recalibrating/retraining [39].

### **SAFE IMPLEMENTATION AND WIDESPREAD USE**

A system-level pathways transformation will be required for artificial intelligence systems to be deployed more broadly and maximize benefits. This in turn would require active participation across all stakeholders in community optometry, hospital-based eye departments, and payers of the system (commissioners), backed by good economic modeling. Such a system-level transformation could help identify the minimal infrastructure requirements/technical capacity that will be needed to make use of artificial intelligence systems/tools, identify mature models or closed systems that could help build the framework for other artificial intelligence systems, and ultimately should be able to demonstrate improvement in patient health outcomes.

Risks/safety implications of artificial intelligence systems could be associated with the underlying technology or the wider system it is implemented in. Methods adopted to monitor safety must be able to detect shortcomings in both. Interpretability of artificial intelligence systems to interrogate the decision-making process is also key and will dictate safety analysis method choice. A recent patient safety analysis [43] of artificial intelligence systems in diabetic retinopathy screening lays the road map for the future depending on the artificial intelligence model. Failure mode and effects analysis [44], which involves creating a detailed map of the process to identify any/all points that could fail, and system-theoretic process analysis, which identifies and analyzes potential safety gaps in systems and their consequence [45], are both amenable to application in artificial intelligence-based screening implementation. Another method called the Bowtie analysis, which is applied in safety-critical industries such as the aviation sector, provides a visual method to identify and map contributing factors to system failure, and solutions to mitigate risk, is now also being applied in health care [46].



Artificial intelligence system implementation in health care requires a financially viable business model for implementation and sustainability. Costs associated with artificial intelligence system adoption could include purchase price, resources, and time required for training, setup, and implementation, including cost of integrating technology into clinical workflow. Sustainability could be achieved if payers and hospitals adopt artificial intelligence models based on anticipated long-term cost implications from improvement in clinical outcome of interest, and reimbursement thresholds are altered to incentivize uptake. For a more comprehensive picture of costs/benefits, other cost efficiencies achieved through artificial intelligence-enabled pathways to patients (reduced hospital visits, travel costs, early treatment) and clinicians (reduced burden to hospital-based departments) should also be considered. To address these issues, future research to evaluate cost-effectiveness of artificial intelligence systems in larger cohorts across multiple artificial intelligence applications in health care and specific-use cases will be important.

There are limited cost–benefit reports from real-world use of artificial intelligence algorithms in clinical practice. Studies in diabetic retinopathy screening using artificial intelligence (semiautomated hybrid approaches) have demonstrated artificial intelligence model usage can be cost-effective [47–50]. A recent health economic assessment in diabetic retinopathy screening using artificial intelligence [43] suggested that where there is no established comparison, as is the case with artificial intelligence systems, cost-utility analysis could be the preferred method, or where clinical outcome needs to be evaluated, cost-effectiveness analysis could be the method of choice. Cost–benefit analysis could be an effective method for intervention assessment in resource-limited settings with high unmet clinical needs.

Despite the vast demonstrable potential of artificial intelligence models, adoption of these systems is still hindered by clinician and patient perception that artificial intelligence is a ‘black box’ [2]. The key to addressing this issue and achieving better human–computer interactions lies in improving artificial intelligence interpretability [2,35] (i.e., explain underlying features that drive decision-making in an understandable way, with emphasis on clinical applicability and effect on patient outcomes). This would help clinicians understand how artificial intelligence systems could assist them with data management, screening, triage, and decision support, enabling them to provide the best possible care to patients, rather than replace them. From the patient perspective, it is important to engage

patients and address concerns that artificial intelligence systems will prevent interaction with clinicians. Interpretable algorithms that are transparent and lead to tangible clinical benefits for patients are more likely to engender trust and be accepted by clinicians and patients alike.

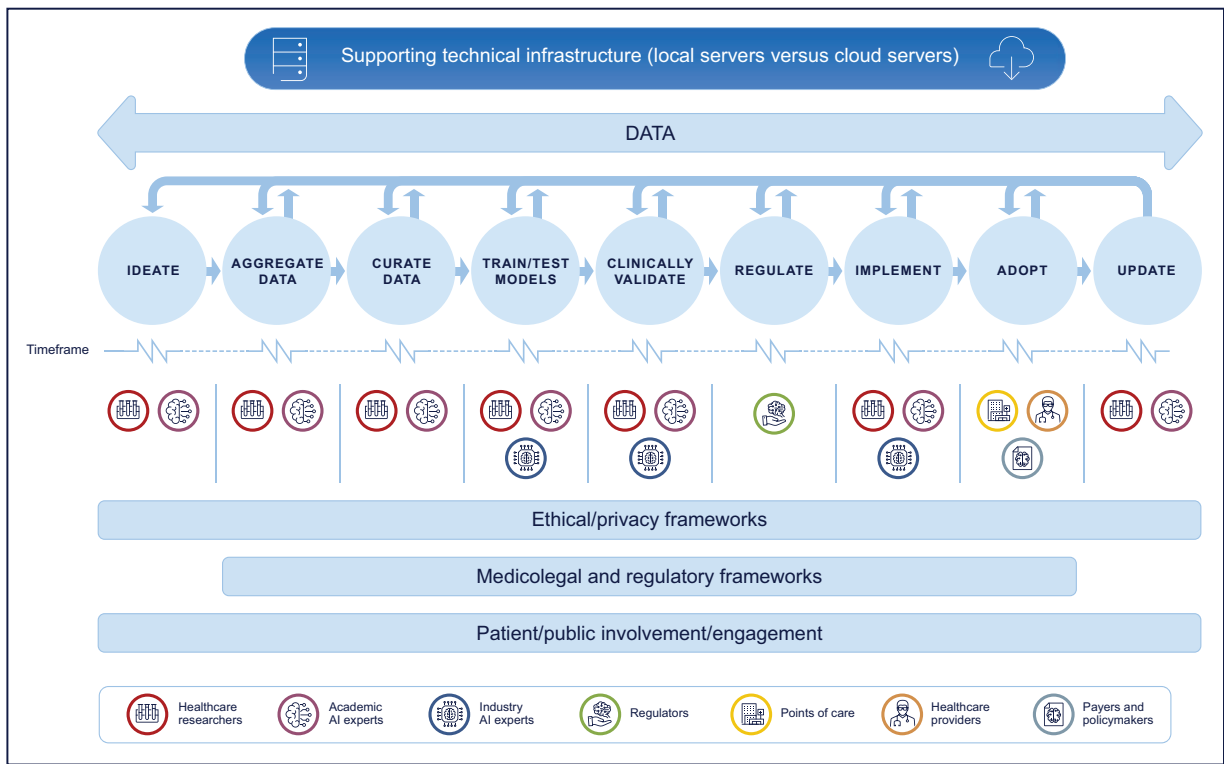
As artificial intelligence enters into medical practice to support clinical decisions, a different set of medicolegal challenges around medical liability, stemming from issues ranging from missed diagnoses to incorrect treatment based on artificial intelligence algorithm recommendations, is a key concern for medical professionals. Laws around medical liability involving artificial intelligence are not developed [51]. A key point is that current laws are dictated by whether standard of care has been adhered to. Because artificial intelligence is new to clinical practice, standards of care around artificial intelligence system usage must be established. Artificial intelligence interpretability will once again be key because it will allow for establishing effective procedures for risk assessment, risk mitigation, oversight, and incident investigation. Understanding around acceptable levels of risk and potential false positives/false negatives generated by an artificial intelligence model will also be key.

Emphasis on any of the above considerations could significantly differ between different countries and resource settings. For example, a retinopathy of prematurity screening tool could have a significant positive impact on healthcare outcomes in systems underresourced for pediatric ophthalmologists, which could affect the approach to risk acceptances. This example highlights the need for implementation science work on a local level adapted to the ‘pain points’ and priority needs of each system.

Trials that assess and validate how such a system-level transformation could be achieved using telemedicine and artificial intelligence; overcome barriers to clinical deployment; address issues of cost-effectiveness and patient and healthcare professional acceptability/trust; perceptions of the ‘black box’ phenomenon; and concerns about impersonal care, training, and familiarization with the technology, are underway (HERMES study) [52]. Work of this kind will be increasingly needed to bridge the ‘artificial intelligence chasm’ and move artificial intelligence systems from the laboratory to front-line clinical practice for the benefit of patients and healthcare systems.

## CALL TO ACTION

To truly deliver PHC in the management of retinal diseases, a system-level paradigm shift enabled by



**FIGURE 1.** Artificial intelligence and personalized medicine in retina: from ideation to adoption.

collaboration of all stakeholders will be needed to tackle existing challenges (Fig. 1). The opportunity to address unmet needs in delivering PHC has been presented by artificial intelligence, and artificial intelligence-enabled telemedicine/virtual clinics [47,53] (e.g., positive OCT picked up by artificial intelligence in the high-street optometry practice triggers a direct appointment to a hospital-based injection clinic, potentially bypassing the intermediate diagnostic step at the hospital, enabling early treatment) are working to demonstrate this promise. However, a commitment to addressing the unmet needs in each stage of development of a PHC solution has to be made.

**Harnessing data**

Clinical healthcare researchers and artificial intelligence experts (academia and industry) must tackle challenges in curation and annotation of large datasets, linking clinical (and other) data to ophthalmic images obtained by various modalities, and standardizing digital imaging and communications in ophthalmology.

**Validating artificial intelligence systems**

Clinical healthcare researchers and artificial intelligence experts (academia and industry) must

collaborate to develop artificial intelligence systems trained on well-curated datasets and real-world data, and develop clinical trials and evidence packages that demonstrate clear benefit and impact of the artificial intelligence system.

**Infrastructure development**

Clinical healthcare researchers and artificial intelligence experts (academia and industry) must work together to develop systems to integrate and manage data, and to support real-time run efforts of deployed artificial intelligence systems.

**Implementation**

Academic and industry experts need to work with policy makers and regulators to develop frameworks for implementation of technically mature artificial intelligence systems into the clinic.

**Adoption**

Healthcare workforce (ophthalmologists/optometrists, clinics, payers, healthcare systems) embrace artificial intelligence systems and create an environment primed for adoption based on demonstrable value of artificial intelligence systems to their patients.

## Patients

Ultimately, patients are at the center of PHC, and it is critical for them to fully engage based on their understanding of artificial intelligence system benefits, in the knowledge that their data and personal health information are protected, and that the ultimate goal of any artificial intelligence system is to improve their treatment journey and outcome.

This is the journey that we as healthcare providers must enable with our patients, from the earliest inception of an artificial intelligence-enabled solution to the impact it will ultimately enable in their care. Ophthalmology is poised to be a leader in the implementation of PHC, and with coordinated efforts across the spectrum of system design and implementation, stands to be an exemplar for other medical specialties.

## Acknowledgements

The authors would like to thank Priya Snehil at Moorfields Eye Hospital, for contributions to the development of the figure, and Anh Duong, Ophthalmology PHC program manager at Roche-Genentech for program and editorial support in development of this review.

## Financial support and sponsorship

Funding was provided by Genentech, Inc. for the study and third-party writing assistance, which was provided by Dinakar Sambandan, PhD, of Envision Pharma Group.

## Conflicts of interest

J.J.H. is an employee of Genentech, Inc. P.A.K. has been a consultant for Apellis, DeepMind, Novartis, and Roche; has received speaker fees from Allergan, Bayer, Heidelberg Engineering, and Topcon; is an equity owner in Big Picture Medical; and receives funding from the Moorfields Eye Charity career development award. K.B. has been a consultant for Novartis and Roche; and has received speaker fees from Alimera, Allergan, Bayer, Heidelberg Engineering, Novartis, and Topcon.

## REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Adamis AP, Brittain CJ, Dandekar A, Hopkins JJ. Building on the success of anti-vascular endothelial growth factor therapy: a vision for the next decade. *Eye* 2020. doi:10.1038/s41433-020-0895-z.
2. Ting DSW, Pasquale LR, Peng L, *et al.* Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 2019; 103:167–175.
3. Ting DSW, Peng L, Varadarajan AV, *et al.* Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res* 2019; 72:100759.
4. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; 25:44–56.  
A comprehensive review article summarizing the available evidence and opportunities for the use of artificial intelligence in medicine.
5. De Fauw J, Ledsam JR, Romera-Paredes B, *et al.* Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018; 24:1342–1350.  
The study demonstrated the clinical applicability of a novel artificial intelligence-enabled framework that analyzes clinical optical coherence tomography scans and makes referral suggestions to a standard that reaches or exceeds that of clinical experts.
6. Raumviboonsuk P, Krause J, Chotcomwongse P, *et al.* Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit Med* 2019; 2:25.
7. Chan S. Crossing the AI chasm [Internet]. Published 2017. Available at: <https://techcrunch.com/2017/01/05/crossing-the-ai-chasm/>. [Accessed 20 May 2020].
8. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. *NPJ Digit Med* 2018; 1:40.
9. Rajkumar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019; 380:1347–1358.
10. Sadda SR, Tuomi LL, Ding B, *et al.* Macular atrophy in the HARBOR study for neovascular age-related macular degeneration. *Ophthalmology* 2018; 125:878–886.
11. UK Biobank. UK Biobank homepage [Internet]. Available at: <https://www.uk-biobank.ac.uk/>. [Accessed 20 May 2020].
12. American Academy of Ophthalmology. IRIS Registry [Internet]. Available at: <https://www.aao.org/iris-registry>. [Accessed 20 May 2020].
13. Parke DW 2nd, Rich WL 3rd, Sommer A, Lum F. The American Academy of Ophthalmology's IRIS<sup>®</sup> Registry (Intelligent Research in Sight Clinical Data): a look back and a look to the future. *Ophthalmology* 2017; 124:1572–1574.
14. INSIGHT. The health data research hub for eye health [Internet]. Available at: <https://www.insight.hdrhub.org/>. [Accessed 20 May 2020].
15. Ada Lovelace Institute. Foundations of fairness: where next for NHS health data partnerships? [Internet]. Published 2020. Available at: <https://understandingpatientdata.org.uk/sites/default/files/2020-03/Foundations%20of%20Fairness%20-%20Summary%20and%20Analysis.pdf>. [Accessed 20 May 2020].
16. Larson DB, Magnus DC, Lungren MP, *et al.* Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. *Radiology* 2020; 295:675–682.
17. The Royal Society. Data governance: public engagement review [Internet]. Published 2017. Available at: <https://royalsociety.org/~media/policy/projects/data-governance/data-governance-public-engagement-review.pdf>. [Accessed 20 May 2020].
18. Notal Vision. Home OCT [Internet]. Available at: <https://notalvision.com/technology/home-oct>. [Accessed 20 May 2020].
19. Keane PA, Topol EJ. Medicine and meteorology: cloud, connectivity, and care. *Lancet* 2020; 395:1334.
20. Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Part of: *Advances in neural information processing systems* 27 (NIPS 2014). <https://papers.nips.cc/paper/5423-generative-adversarial-nets>. [Accessed 20 May 2020].
21. Ravuri S, Vinyals O. Classification accuracy score for conditional generative models [arXiv:1905.10887] [Internet]. Published 2019. Available at: <https://arxiv.org/abs/1905.10887>. [Accessed 20 May 2020].
22. Komorowski M, Celi LA, Badawi O, *et al.* The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018; 24:1716–1720.
23. Liu Y, Logan B, Liu N, *et al.* Deep reinforcement learning for dynamic treatment regimes on medical registry data. *Healthc Inform* 2017; 2017:380–385.
24. Faes L, Wagner SK, Fu DJ, *et al.* Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. *Lancet Digit Health* 2019; 1:e232–e242.  
The study demonstrated the design and implementation of an artificial intelligence enabled in medical image classification system by healthcare professionals with no artificial intelligence experts, potentially overcoming the need for specialized artificial intelligence expertise in the development of artificial intelligence models.
25. US Food and Drug Administration. Software as a medical device (SaMD) clinical evaluation. Guidance for Industry and Food and Drug Administration staff [Internet]. Published 2017. Available at: <https://www.fda.gov/media/100714/download>. [Accessed 20 May 2020].
26. US Food and Drug Administration. Software as a medical device (SaMD) [Internet]. Published 2018. Available at: <https://www.fda.gov/medical-devices/digital-health/software-medical-device-samd>. [Accessed 20 May 2020].
27. Long E, Lin H, Liu Z, *et al.* An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nat Biomed Eng* 2017; 1:0024.
28. Milea D, Najjar RP, Jiang Z, *et al.* Artificial intelligence to detect papilledema from ocular fundus photographs. *N Engl J Med* 2020; 382:1687–1695.
29. Gulshan V, Peng L, Coram M, *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; 316:2402–2410.
30. Kiani A, Uyumazturk B, Rajpurkar P, *et al.* Impact of a deep learning assistant on the histopathologic classification of liver cancer. *NPJ Digit Med* 2020; 3:23.

31. Abràmoff 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. *NPJ Digit Med* 2018; 1:39.
- A pivotal prospective trial that demonstrated the ability of an artificial intelligence-enabled system to detect diabetic retinopathy, and supported the US Food and Drug Administration authorization of the first artificial intelligence diagnostic system for use by healthcare providers.
32. Bellemo V, Lim ZW, Lim G, *et al*. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digit Health* 2019; 1:e35–e44.
33. Kanagasingam Y, Xiao D, Vignarajan J, *et al*. Evaluation of artificial intelligence-based grading of diabetic retinopathy in primary care. *JAMA Netw Open* 2018; 1:e182665.
34. Liu X, Faes L, Kale AU, *et al*. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health* 2019; 1:e271–e297.
35. Kelly CJ, Karthikesalingam A, Suleyman M, *et al*. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; 17:195.
36. European Union (EU). Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 [Internet]. Published 2017. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0745>. [Accessed 20 May 2020].
37. Health Education England. The Topol review: preparing the healthcare workforce to deliver the digital future [Internet]. Published 2019. Available at: <https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf>. [Accessed 20 May 2020].
38. US Food and Drug Administration. Clinical decision support software: draft guidance for industry and FDA staff [Internet]. Published 2019. Available at: <https://www.fda.gov/media/109618/download>. [Accessed 20 May 2020].
39. US Food and Drug Administration. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) [Internet]. Published 2019. Available at: <https://www.regulations.gov/document?D=FDA-2019-N-1185-0001>. [Accessed 20 May 2020].
40. Shah P, Kendall F, Khozin S, *et al*. Artificial intelligence and machine learning in clinical development: a translational perspective. *NPJ Digit Med* 2019; 2:69.
41. BSI Group America Inc. The postmarket priority. Understanding and meeting demand for effective postmarket clinical follow-up [Internet]. Published 2013. Available at: <https://www.bsigroup.com/globalassets/meddev/localfiles/de-de/whitepapers/bsi-md-the-postmarket-priority-whitepaper-de-de.pdf>. [Accessed 20 May 2020].
42. Tariah I, Pine R. Effective postmarket surveillance. Understanding and conducting vigilance and postmarket clinical follow-up [Internet]. Available at: <https://www.bsigroup.com/meddev/LocalFiles/en-US/Whitepapers/WP-Postmarket-surveillance.pdf>. [Accessed 20 May 2020].
43. Xie Y, Gunasekeran DV, Balaskas K, *et al*. Health economic and safety considerations for artificial intelligence applications in diabetic retinopathy screening. *Transl Vis Sci Technol* 2020; 9:22.
44. Robinson DL, Heigham M, Clark J. Using failure mode and effects analysis for safe administration of chemotherapy to hospitalized children with cancer. *Jt Comm J Qual Patient Saf* 2006; 32:161–166.
45. Leveson NG, Thomas JP. STPA handbook [Internet]. Published 2018. Available at: [http://psas.scripts.mit.edu/home/get\\_file.php?name=STPA\\_handbook.pdf](http://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf). [Accessed 20 May 2020].
46. McLeod RW, Bowie P. Bowtie analysis as a prospective risk assessment technique in primary healthcare. *Pol Pract Health Saf* 2018; 16:177–193.
47. Kern C, Fu DJ, Kortuem K, *et al*. Implementation of a cloud-based referral platform in ophthalmology: making telemedicine services a reality in eye care. *Br J Ophthalmol* 2020; 104:312–317.
48. Scotland GS, McNamee P, Philip S, *et al*. Cost-effectiveness of implementing automated grading within the National Screening Programme for diabetic retinopathy in Scotland. *Br J Ophthalmol* 2007; 91:1518–1523.
49. Tufail A, Kapetanakis VV, Salas-Vega S, *et al*. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. *Health Technol Assess* 2016; 20:1–72.
50. Xie Y, Nguyen Q, Bellemo V, *et al*. Cost-effectiveness analysis of an artificial intelligence-assisted deep learning system implemented in the national telemedicine diabetic retinopathy screening in Singapore. *Invest Ophthalmol Vis Sci* 2019; 60:5471.
51. Price WN II, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. *JAMA* 2019; 322:1765–1766.
52. National Institute for Health Research. Research award. Tele-opHthalmology-enabled and ARtificial Intelligence-ready referral pathway for coMmunity optomEtry referralS of retinal disease: the HERMES study – a Cluster Randomised Superiority Trial with a linked Observational Diagnostic Accuracy Study [Internet]. Available at: <https://fundingawards.nihr.ac.uk/award/NIHR127773>. [Accessed 20 May 2020].
53. Korot E, Wood E, Weiner A, *et al*. A renaissance of teleophthalmology through artificial intelligence. *Eye* 2019; 33:861–863.