Integrating artificial intelligence into an ophthalmologist's workflow: obstacles and opportunities

Priyal Taribagil, HD Jeffry Hogg, Konstantinos Balaskas & Pearse A Keane

Published online: 07 Feb 2023.

Submit your article to this journal ☑

Article views: 13

View related articles ☑

View Crossmark data ☑
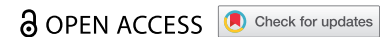
Taylor & Francis
Taylor & Francis Group

REVIEW

🔓 OPEN ACCESS ✅ Check for updates

Integrating artificial intelligence into an ophthalmologist's workflow: obstacles and opportunities

Priyal Taribagil[a], HD Jeffry Hogg[a,b,c], Konstantinos Balaskas[d,e] and Pearse A Keane[d,e]

[a]Medical Retina Department, Moorfields Eye Hospital NHS Foundation Trust, London, UK; [b]Department of Population Health Science, Population Health Science Institute, Newcastle University, Newcastle upon Tyne, UK; [c]Department of Ophthalmology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Freeman Road, Newcastle upon Tyne, UK; [d]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust, London, UK; [e]Medical Retina, Institute of Ophthalmology, University College of London Institute of Ophthalmology, London, UK

ABSTRACT

Introduction: Demand in clinical services within the field of ophthalmology is predicted to rise over the future years. Artificial intelligence, in particular, machine learning-based systems, have demonstrated significant potential in optimizing medical diagnostics, predictive analysis, and management of clinical conditions. Ophthalmology has been at the forefront of this digital revolution, setting precedents for integration of these systems into clinical workflows.

Areas covered: This review discusses integration of machine learning tools within ophthalmology clinical practices. We discuss key issues around ethical consideration, regulation, and clinical governance. We also highlight challenges associated with clinical adoption, sustainability, and discuss the importance of interoperability.

Expert opinion: Clinical integration is considered one of the most challenging stages within the implementation process. Successful integration necessitates a collaborative approach from multiple stakeholders around a structured governance framework, with emphasis on standardization across healthcare providers and equipment and software developers.

1. Introduction

Ophthalmology is one of the busiest outpatient specialties in the UK and forms a major component of healthcare worldwide. The National Healthcare System (NHS) is one of the largest healthcare providers in the world and noted to have over 7.9 million ophthalmology clinic attendances per year [1]. Demand in healthcare workforce and clinical services is predicted to rise by a further 30–40% over the next 20 years [2]. Contributing factors include aging population, increased prevalence of complex, chronic conditions, and added pressure on timely detection and diagnosis. Capacity pressures have significantly worsened since the COVID-19 pandemic [3], further emphasizing the clinical need for a more efficient system. Artificial intelligence (AI) has the potential to help address these challenges, by augmenting or automating various processes across the healthcare ecosystem. Integration of AI platforms and systems into clinical workflows provides opportunities to increase clinical efficiency and productivity, enabling the consistent delivery of higher quality care despite growing demand.

Artificial intelligence uses the support of technology to simulate elements of human intelligence, cognitive processes and functional decision-making [4]. The term was first coined by McCarthy in 1956 and since then has emerged as an important research field with numerous potential benefits to healthcare. AI continues to have varied definitions, but is commonly described as a machine-based system that has

the potential to produce certain desired outputs, predictions, and decisions [5]. Within healthcare, it can still be taken to include computerized decision support tools, which automate the delivery rules directly dictated by human experts, in rule-based decision trees[6]. Machine learning is a subset of AI where systems formulate associations based on previous data, with potential for simultaneous performance optimization. It is machine learning-enabled clinical tools that have sparked the ongoing surge in government, academic, and industry interest and that will form the focus of this review.

There are many machine learning algorithms used in healthcare such as linear regression, logistic regression, decision trees, and random forests. None of these examples represent deep learning, which is a subset of machine learning that uses multiple layers of nodes connected in a neural network. These neural networks are able to process multiple data items, whilst preserving spatial distribution [7]. The incorporation of hidden layers aids the exploration of more complex non-linear data patterns [8]. Convolutional Neural Networks (CNN) are a subtype of neural networks used commonly in image recognition. By using multiple convolutional layers, they are able to process both simple and complex features (edges, lines, colors, shapes, etc.). Current examples of CNNs include AlexNet, GoogleNet, and ResNet[9]. Much of the successes in deep learning have been driven by CNNs.

Before delving into the implementation process of these machine learning-enabled tools, it is helpful to have an

Article highlights

- Clinical AI has sparked ongoing academic, industry, and government interest. Applications to healthcare, have the potential to optimize clinical diagnostics and predictive analytics.
- Ophthalmology is a leading specialty in the integration of AI in clinical workflows, due to the abundance of clinical imaging data and the value of machine learning-enabled tools for their interpretation.
- Incorporating machine learning-enabled tools into real-world ophthalmology workflows is challenging. Input from a variety of stakeholders (clinicians, developers, research groups and overseeing government and regulatory bodies) is vital to maximize adoption.
- CONSORT AI, STARD-AI, TRIPOD-AI and DECIDE-AI are standards that should guide reporting and evaluation of clinical AI, promoting transparency.
- Limited explainability and interoperability with established health information systems are key barriers in the integration process.
- Integration of AI into clinical workflows should be conducted in concordance with appropriate ethical and legal standards. Accountability for machine learning systems is challenging, but a pathway should be created for clinical incident investigation and in cases of medical malpractice.

understanding of common learning methods used within machine learning. The three main types include supervised, unsupervised, and reinforcement learning. In supervised learning, a model is trained using already labeled data; with input weights used to formulate model predictions and outputs. Unsupervised learning involves training a model with unlabeled data, allowing associations to be made between datapoints purely based on pattern recognition [9]. From previous literature, most of the models that are used in clinical decision-making settings are supervised machine learning models [10]. In particular, there have been recent drive toward self-supervised learning. The first stage of this involves training on unlabeled data, in the form of a preparation task, 'pretext task'[11]. The subsequent stage involves training on a labeled dataset with the aim of applying the knowledge learnt to specific medical tasks. Reinforcement learning is the third main type of deep learning and is based on the concept of intelligent agents taking actions in order to maximize cumulative reward. Reinforcement learning has demonstrated impressive applications in gaming. Its application in healthcare, though limited at present, is a promising area of intense academic activity[10].

Substantial progress of deep learning tools has been noted in a number of medical disciplines including radiology [12–14], respiratory medicine [15], cardiology [16], and dermatology [17]. Use cases range from the identification of pneumonia on chest x rays [18] to the classification of malignant skin lesions [17,19] with the use of CNNs. Ophthalmology has been at the forefront of this digital revolution with significant contributions made within the field of image recognition [20]. This is unsurprising given the abundance of imaging data that is generated on a daily basis to inform patient management [21]. Appropriate integration of machine learning into the ophthalmology workflow has the potential to aid clinical diagnostics, disease grading, management regimens and improve patient monitoring. A further benefit of deep learning is the optimization of screening programs, e.g. in retinopathy of

prematurity and diabetic retinopathy screening [22]. Additionally, AI is used to deliver more personalized healthcare by optimizing patient-specific treatment, catering to the natural heterogeneity present within the general population [23]. Despite all of these pre-clinical advances, incorporating machine learning-enabled tools into real-world ophthalmology workflows is a challenging prospect.

This review will delve into the interdependent factors that will shape AI integration in ophthalmology while discussing the key concepts of ethics and regulation, clinical governance and potential barriers to clinical adoption.

2. AI in ophthalmology

Within the field of ophthalmology, AI research has been applied to many conditions including but not limited to diabetic retinopathy [24,25], age-related macular degeneration (AMD) [25–28], glaucoma [29] and cataracts [9,30]. As of November 2022, there were over 1000 studies of ophthalmic applications of AI listed in PubMed [31,32]. A more comprehensive discussion of individual studies is beyond the scope of this review and has been recently addressed elsewhere [22]. Despite the substantial volume of AI research in ophthalmology, there are very few studies that report the implementation of machine learning-enabled tools in real-world settings. We wish to highlight this gap and discuss some key examples within current literature and clinical practice.

The UK-based National Institute of Health Research (NIHR) HERMES study (teleophthalmology-enabled and artificial intelligence-ready referral pathway for community optometry referrals of retinal disease) is a multicentered implementation study based on the interaction between tele-medicine and AI decision support tool in referral pathways. The diagnostic accuracy, cost-effectiveness, and performance of the machine learning-enabled clinical decision support (CDS) will be assessed prospectively [33]. This is one of the first large-scale implementation studies within the field. At the time of writing, only one qualitative study of a machine learning-enabled ophthalmic workflow has been published and three other qualitative studies have been conducted focusing on patient and clinician perspectives of hypothetical machine learning-enabled ophthalmic workflows [34–37]. Most of these publications are concerned specifically with machine learning-enabled diabetic retinopathy screening, but cover diverse settings across New Zealand, Thailand, Germany, and the USA. Across these studies it is apparent that at a high level there is a strong value proposition for clinical AI in ophthalmology and wide acceptance. However, closer examinations of specific use cases surface complex interdependent factors that influence implementation outcomes. It is only by further developing this small qualitative literature that the wealth of efficacy-focused quantitative research in ophthalmic AI can be efficiently implemented into real-world workflows [38].

In the USA, there are a number of United States Food and Drug Administration (FDA)-approved AI systems currently in practice. IDx-DR was one of the first fully autonomous AI-based diabetic retinopathy diagnosis systems to be approved by the FDA and integrated into primary care. The results of the original study demonstrated a sensitivity of 87.2% and a

specificity of 90.7% for identifying mild diabetic retinopathy and/or diabetic macular edema equivalent in the Early Treatment Diabetic Retinopathy Study (ETDRS) scale [39]. The system had a 97.6% sensitivity for identifying individuals requiring urgent assessment according to American Academy of Ophthalmology Preferred Practice Pattern (PPP) [39]. Since its approval, IDx-DR has been used at John Hopkins, Stanford and Mayo Clinic for diabetic retinopathy screening [40]. EyeArt (Eyeuk) is another autonomous, cloud-based, FDA approved system used in the US for identifying and grading diabetic retinopathy [40]. Collaborative work between Google Health and Optos has resulted in the availability of machine learning-enabled tools for ultra-widefield imaging. This has potential to aid rapid screening and identification of diabetic retinopathy and macular edema [41]. Other clinical decision support tools available for OCT systems include RetInSight, used for monitoring of patients with neovascular age-related macular degeneration [42,43]. RetinAI also provides an AI model enabling further quantification of retinal diseases, with wider contributions to optimizing the overall healthcare ecosystem [44]. Both these tools have met the 'Medical Device Directive' regulatory requirements as part of the European Union (EU) legislative framework [42,44]. RetinAI has also recently received clearance from the FDA for their medical imaging ophthalmology platform[44].

3. Integration of AI systems in ophthalmic workflows

The machine learning pipeline involves collecting and pre-processing data, training the model, validating and ultimately deploying an AI system into a clinical pathway. Despite there being significant advances in AI systems, on a smaller individual level, there is still a significant gap between development and implementation into clinical practice. This bridge, commonly known as the 'AI Chasm' is a rate-limiting step to widespread integration [45]. Despite the widely perceived potential of AI to improve patient outcomes and streamline clinical workflows, care must be taken to ensure that the implementation process is carried out with satisfactory

regulatory compliance. It is also pertinent that any large-scale integration adheres to a fixed AI governance structure thereby ensuring consistency across clinical institutions.

Implementation is a multifaceted concept, with clinical integration often considered as the hardest stage[10,46]. Optimizing this requires input from multiple stakeholders with continual transparency throughout the process. The stakeholder ecosystem includes clinicians, developers, research groups, and overseeing government and regulatory bodies [47].

3.1. Translational pathway

The following section provides a brief summary on the translational pathway, with a particular emphasis on how integration fits into the overall process (Figure 1).

The first stage in the AI system lifecycle involves designing and developing a deep learning AI-based system or tool, either in-house or via an external developing company. Sources of funding include university, industry, or government, to name a few. Careful consideration should be given to select a clinically meaningful use case that helps to overcome gaps within the current health systems. Although this seems to be rudimentary, it is important to be able to distinguish use cases that would benefit from machine learning-enabled tools rather than clinical pathway redesign or other digital health solutions. At this initial stage, there is also the need for development of large representative datasets, on which subsequent training and validation can take place. Publicly available ophthalmology imaging datasets (e.g. Messidor, Kaggle etc.) can be useful resources during the initial designing stages [48].

The next stage involves formal evaluation and validation. Analysis of performance metrics such as accuracy and reliability helps to assess the scientific robustness of the AI product and determine its statistical validity. It is also important to review the clinical and economic utility before proceeding with implementation. Validation on external datasets helps to assess these features, with potential for fine-tuning of the tool. In addition to reaching the required technical
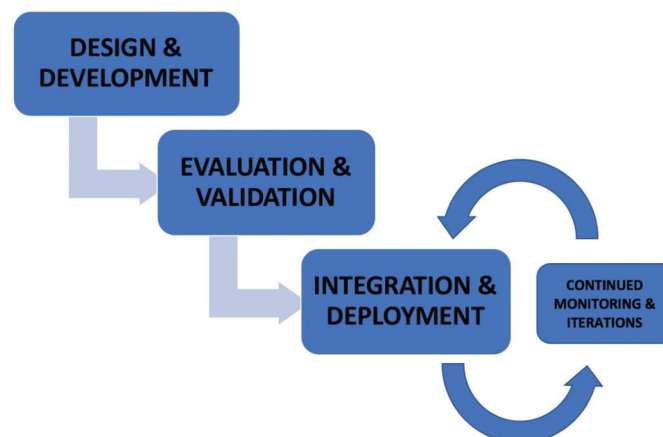

Figure 1. Diagram summarizing how integration relates to other components of the implementation process.

performance measures, the ease of product–user interaction should be evaluated, with an additional focus on safety, liability, safeguarding, and ethical integrity—topics that will be covered in this review.

Clinical integration is a challenging step within the implementation pipeline, as sustainable incorporation of the AI product into existing clinical workflows does not come without its difficulties. Previous studies have identified two modes that are commonly used in AI governance frameworks: 'shadow' and 'canary' clinical deployment. In 'shadow' mode, there is real-time deployment of the AI model without any clinical implications. The AI system acts in the background (in parallel) to the current existing clinical workflows. Iterations are made subsequent to its performance in pre-selected metrics. The 'canary' mode involves a graded introduction of the model into clinical settings with active regulation and feedback mechanisms in place [49,50].

Ongoing surveillance and regulation will be required post-deployment to ensure optimal effectiveness, safety and to prevent prediction drift [51]. A prerequisite for this is the building of suitable infrastructure that can cope with maintaining, regulating, and storing the data flow. Comparison of the AI-guided diagnosis to the clinician's recommendations can be used to assess the accuracy of the product. Studies have identified the positive role of MLOps (machine learning operations) in monitoring post-deployment, in non-healthcare-based use cases [49,52]. Similar applications could be considered within healthcare, particularly within ophthalmology.

4. How could AI be implemented into ophthalmology workflow

Table 1 summarizes the different components of an ophthalmology workflow and how AI can be integrated into this. In terms of workflow, there is often an opportunity to implement tools to assist with or automate clinical decisions. This often has significant implications for the value an AI tool can offer in terms of healthcare cost, quality, and efficiency. It is also, rarely the case that such a clear dichotomy between automation and assistance exists in practice as AI outputs intended to assist end-users can often be rejected out of hand or accepted without critique, depending on various adopters, use case and contextual factors [53,54].

5. Challenges

As discussed thus far, we have shown how artificial intelligence has transformative potential for clinical ophthalmology. However, implementation of these clinical tools in real-world settings will require us to overcome some major challenges, all of which contribute to the 'AI Chasm." These barriers can be represented within three categories of clinical, technical, and operational although some issues may span more than one of these groups. A distinguishing feature of clinical AI among other digital health technologies is that it requires a great depth and breadth to address these challenges. From a clinical perspective, technology intrudes much deeper into the clinical decision-making process, requiring close guidance and evaluation from clinicians. This is similarly reflected from a technical perspective, where not only the development of models but also the anticipation and mitigation of drifts in performance across time or sites demand the attention of skilled data scientists and engineers. Finally, from an operational point of view, the challenges posed by a rapidly evolving regulatory landscape alongside the uncertainties associated with this novel technology are also highly demanding. We examine these in challenges using the above framework to provide a structured approach to addressing these issues.

Table 1. Highlighting different ways AI could be used within ophthalmology workflows.

Output	Clinical use
Diagnosis	This can be considered in two ways: • Assistive: AI system supports the clinician by indicating abnormalities with the scan • Autonomous: AI system evaluates the scan, stating the presence or absence of the disease without the involvement of a clinician. Use case: Presence of glaucomatous changes to the optic disc based on fundus and/or OCT images. Other examples include diabetic retinopathy grading, ROP severity grading or presence of plus, pre-plus disease.
Screening	Analyze image or tabular data to highlight individuals that are at rick of developing a disease [59]. This can be considered in two ways: • Assistive: AI system flags up potential high-risk individuals, to make the clinician aware prior to clinical assessment. This can help direct the consultation appropriately. • Autonomous: AI system is incorporated into the screening process, only referring high-risk patients for further clinical assessment. Use case: Hydroxychloroquine toxicity (based on OCT image), diabetic retinopathy screenings.
Monitoring	This could be considered in two ways: • Assistive: AI system supports the clinician in surveillance of clinical conditions. • Autonomous: AI system outputs the severity and correlates with the need for retreatment, without clinician input. Use case: Guidance regarding anti-VEGF retreatment [59].
Predication	Predictive modelling can also be used on electronic patient records to predict the natural progression of the disease [59]. Use case: AI used to predict and individual's conversion to late or severe disease. This could guide decisions regarding follow-up intervals and treatment options.

AI – Artificial Intelligence, ROP – retinopathy of prematurity, OCT – optical coherence tomography, anti-VEG-F – anti-vascular endothelial growth factor

5.1. Clinical challenges

5.1.1. Clinical validation

Before integration of clinical tools in real-life workflows, robust evaluation is required with particular emphasis on clinical performance and safety. There are several valuable guidelines that are available globally for evaluating and reporting the results of an AI model. Consolidation Standards of Reporting Trials (CONSORT-AI) is an extension of CONSORT 2010 reporting guidelines and is used to evaluate clinical trials involving AI. This international initiative is aimed at promoting transparency, enabling a rigorous review of design, checking the quality of input data, data handling, and evaluation of performance errors [55,56]. Standards for Reporting of Diagnostic Accuracy Studies (STARD-AI) and Transparent Reporting of a Multivariable Prediction Model of Individual Prognosis or Diagnosis (TRIPOD-AI) can be used for evaluation of diagnostic accuracy or prognostic model studies [57]. These reporting guidelines provide a clear and rigorously constructed definition of what transparent reporting of different stages of clinical AI research involves. New reporting guidelines, Developmental and Exploratory Clinical Investigation of Decision-support systems driven by Artificial Intelligence (DECIDE-AI) have been designed to help bridge the development to implementation gap [58]. It aims to provide reporting guidance on early clinical evaluation of decision support systems [59]. Although all the guidelines mentioned above provide excellent reporting advice, they largely target the academic and regulatory community and may not be so accessible to end-users and local staff responsible for the procurement of clinical AI. Clear, succinct communication about the machine learning model should be made available, to aid successful integration into clinical workflows. This need has also begun to be addressed through complementary documents with wider accessibility [60]. Algorithmic impact assessments are another such method designed to promote accountability for the designing and integration of AI systems. This involves assessment of societal impacts and long-term consequences prior to widespread integration and can be used in conjunction with other accountability measures [61].

5.1.2. Need for explainability & interpretability

Interpretability is the "extent to which an observer understands the cause of a decision" [62]. This is based on the design of the model and is not representative of the accuracy of the outputs. Explainability delves into this further by attempting to understand the reasons for the association by focusing on the internal mechanics of the machine learning system [63]. At present, there is often a trade-off between performance and explainability. Research has shown that in certain use cases, models (e.g. deep learning) with greater performance are often least explainable. On the contrary, those with poorer performance (e.g. decision trees & linear regression) often provide more explainability. Therefore, it is more appropriate to consider them on a spectrum and analyze models based on the degree to which they fit these concepts [64]. This helps us to decide which class of machine learning algorithms are most appropriate to employ for specific use

cases and be mindful of the compromises regarding explainability (Figure 2).

Reluctance in implementation of AI into clinical workflows is partly governed by the issue of limited model explainability. At present, deep learning models are considered as "black boxes." This principle in computer science refers to a concept of being unable to directly view or understand the connections formed between the input and output [65]. As neural networks consist of multiple hidden layers, their complex connections and pattern recognition may lack robust explainability. From the medical perspective, understanding how a deep learning model reaches the output value can help to promote trust, confidence, and transparency in the algorithm. Explainable AI may help facilitate faster clinician acceptance [66]. In cases of disagreement, explainability can also help clinicians to make informed decisions about whether to adhere with the system's recommendations, something referred to as 'AI interrogation practices' by Lebovitz et al [53,67]. From a patient perspective, providing explainable diagnostics makes them feel more informed and empowered, ultimately supporting shared decision-making [67]. In addition to the clinical barriers discussed, explainable AI may also help to mitigate potential ethical, legal, and regulatory concerns [67]. Hence, as a research community, we need to explore this concept further, as it may help to produce effective models and promote an environment of trustworthy artificial intelligence, which is particularly prudent in healthcare settings.

5.1.3. Human-computer interaction and issues of clinician adoption

Human-computer interaction (HCI) is defined as 'designing, evaluating, and using information & communication technologies' to 'improve user experience, task performance, and overall quality of life'[68]. This is being taken into consideration when developing AI systems as design and usability will affect the success of a system through its influence on user adoption. Clinician adoption is another related factor that has a great influence over AI implementation. Despite the existence of clinically validated AI models with corresponding supportive business strategies, difficulties may still arise in gaining user acceptance [69]. This adds another level of complexity in implementation of CDS platforms and tools in the real-world setting. Collaborative research and engagement from a wider range of stakeholders is essential to maximize usability [70].

Singer et al. have identified four phases in the development and implementation process – which incorporate input from a variety of stakeholders [71]. This can be described as below:

- **Phase 1 – Iterative co-identification**: This involves identification of the initial user needs and development of the initial machine learning tool, e.g. designing a risk stratification tool for individuals and embedding this into EHR.
- **Phase 2 – Iterative co-engagement**: This involves collaboration and engagement with multiple stakeholders including users, developers, and external experts. This results in adjustment of the machine learning tool.
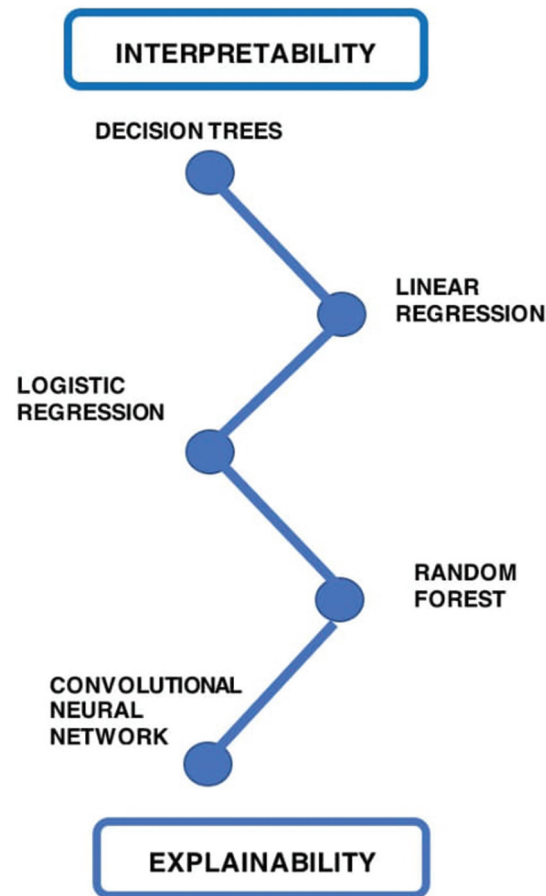
Figure 2. Diagram highlighting the spectrum of interpretability & explainability.

- **Phase 3 – Iterative co-application**: During this, users identify discrepancies in the data, with the developer aiming to reconcile data across various groups.
- **Phase 4 – Iterative co-refinement**: This involves further analysis and assessment of inaccuracies, enabling further adjustment of the machine learning tool.

The importance of input from multiple stakeholders [71] is demonstrated by the model. This provides opportunities for context-specific, interdisciplinary co-operation as user preferences can be integrated throughout the design process.

Furthermore, recruitment of individuals to the pathway adds a degree of human communication and interaction to the output of the machine learning model [72]. It also has a potential to minimize the education gap across the wider workforce. An example can be taken from the integration of machine learning sepsis early warning system into an emergency medicine clinical workflow [73]. Here, input from rapid response team, operating on the recommendation of the machine learning algorithm, provided an intermediate step in the AI implementation pipeline [73]. Similarly in a primary care screening tool for peripheral arterial disease, a dedicated secondary care multi-disciplinary team handled all of the AI outputs and communicated decisions through conventional channels to primary care end-users [74]. Consideration of

human-computer interaction can help to address some of the barriers to clinicians adopting AI in clinical workflow [75].

Resistance to clinical adoption can be further addressed by the launch of virtuous adoption cycles where the AI tool directly benefits the adopter or rewards them. This could be achieved by making the associated administration processes easier for users [76]. Such an understanding of what adopters value and how to deliver that value directly or indirectly through implementation of an AI-enabled tool can only be achieved through meaningful multi-stakeholder engagement.

5.2. Technical challenges

5.2.1. Integration with electronic health records

Further to the integration effort, sufficient infrastructure is required to manage and process large data, often held within electronic health records (EHR). These contain multi-modal data that add a degree of complexity to the integration process. Current EHR systems are inflexible, not user-friendly and are often incongruous with other clinical platforms. Furthermore, there is no common data model available for EHR data within ophthalmology. Application of the 'Fast Healthcare Interoperability Resource' (FHIR) to ophthalmic healthcare has the potential to optimize health data interoperability [77,78]. This includes publishing an implementation

guide that provides international guidance regarding standardized clinical terminologies.

Unfortunately, there are limited incentives for EHR providers to address these issues and it lends a market advantage to AI tools, which they develop, and embeds within their EHR products, which risk adoption for convenience rather than performance [79]. AI could be an additional incentive to make these more streamlined, patient-centered and adaptable. In particular, it can encourage policymakers and providers to demand greater interoperability of EHR and equipment vendors

Machine learning models can be applied to data from EHR in an autonomous manner or as clinical decision support (CDS). Despite the regulatory and ethical considerations, automated processes demonstrate increased efficiency. Machine learning enabled CDS work in conjunction with clinicians, improving the overall delivery of care by providing diagnostic assistance, thereby supporting their existing evidence-based practice. Currently, there is no gold standard for CDS model structure; however, the two common frameworks recognized in current literature include centralized and decentralized workflows [52]. In a centralized workflow, the user of the CDS is not in direct contact with the patient and hence all principal decisions are made by an overseeing command center. On the contrary, decentralized workflows involve direct interaction between the CDS and the user/patient interface. These models are easier to implement and may not require complete or widely scaled integration into an EHR [10,52,80].

Furthermore, considerations should be given regarding local versus cloud deployment of models. Local deployment is dependent on healthcare institutions having appropriate IT infrastructure, with independent privacy and security regulations. Disadvantages of 'on-premise' software include problems maintaining complex programs and models [81]. Cloud software models demonstrate superior scalability as evaluation of imaging data is performed and stored in a remote cloud-based platform [82]. Many current AI systems such as RetinAI and RetinSight are cloud-based. Rigorous data protection policies and regulations are hence required in such deployment. Questions also arise regarding the feasibility of integration in busy, high-volume ophthalmology clinics, given the possible time delay to obtain an output, as these systems are dependent on variable factors such as internet connection.

5.2.2. The need for interoperability with use of DICOM standards

Data relevant to clinical decision-making in ophthalmology are often fragmented across different platforms and systems creating interoperability challenges for AI-enabled tools. Data interoperability is fundamental for successful AI integration as it provides shared and equal access to data resources. Multiple imaging modalities (optical coherence tomography – OCT, widefield fundus imaging, etc.) are involved in ophthalmic clinical practice. This prioritizes the need for common standards to facilitate the research and application of ophthalmic AI-enabled tools [83]. Variations that exist between imaging equipment, software, and data storage formats hinder data interoperability, limiting the ease with which multi-modal data

can be assimilated and analyzed. Digital Imaging and Communications in Medicine (DICOM) is the international standard for medical imaging used to view, store, retrieve, and share images [84]. It was created with the intention of maintaining standards and consistency across the varying imaging modalities.

Radiology was one of the first fields to embrace the challenge and incorporate the DICOM guidelines due to the specialty's dependence on multi-modal imaging data. In ophthalmology, compliance to these guidelines remains variable globally with smaller individual research groups developing their own image acquisition and analysis protocols, dictated by the restrictions posed by their local equipment and software [85]. These approaches may not be scalable and pose difficulties in replicability [66]. In 2021, the American Academy of Ophthalmology (AAO) published a report encouraging manufacturers to comply with the DICOM standards. The Royal College of Ophthalmology, Asia-Pacific Academy of Ophthalmology, and Royal and Australian and New Zealand College of Ophthalmology all supported this initiative [86,87]. DICOM compatibility is an important goal; however at times, data may be only 'partially' DICOM-compatible and still pose interoperability challenges [88]. Global consensus among vendors and providers will be required to maximize interoperability, ultimately enabling AI to provide efficient clinical and diagnostic support to system users.

5.2.3. Surveillance and monitoring

The next and arguably most crucial barrier to sustainable and successful implementation is ensuring that there is ongoing performance monitoring of the machine learning-enabled tool. This is essential from a safety point of view due to the challenging nature of errors in AI-based systems, which make them hard to identify. There are numerous reasons why machine learning-enabled tools may not perform optimally ranging from poor data quality to external factors such as adversarial attacks, which can compromise the algorithmic outcome [51]. Although effort to minimize these factors should be made throughout the AI pipeline, it is important to acknowledge that at times, they may only emerge after integration. Furthermore, despite the adaptive nature of an AI system, risk of 'data drift' can contribute to poor clinical performance of a tool [89]. This is particularly relevant with the emergence of a new disease (e.g. COVID) or change in diagnostic classification for a clinical condition. Continuous monitoring of accuracy and performance is essential to monitor for data drift and provides developers with scope to retrain the model, should this be required.

Providing a uniform framework to aid monitoring can help to identify and pro-actively solve problems and mitigate errors as they appear throughout the implementation process. This can be further facilitated with a comprehensive risk assessment process similar to that conducted by manufacturers when deploying medical devices into real-world settings. Regardless of which monitoring framework is adopted, appropriate infrastructure is also required to store the data, maintain data flow, and provide automated regulation [90].

The 'medical algorithmic audit' exemplifies such a framework, aiming to monitor the ongoing performance of an AI-based tool and investigate and pre-empt potential sources of error that may be generated throughout the implementation process [51]. Liu et al. have adapted this framework for use in medical AI-enabled tools such that developers have the opportunity to modify the artificial intelligence system, addressing the issues identified by the audit process before and after real-world integration [51]. This algorithmic audit has been applied to evaluate a hip fracture machine learning-enabled tool [51].

The International Organization for Standardization (ISO) also provides standards on risk management, performance evaluation, and statutory requirements of medical devices and their associated software. Specific guidance is now available (ISO 34791) on evaluating medical technology that uses machine learning and artificial intelligence [91]. This includes consideration of potential hazards, technical and security-based features of the tool, and evaluation of the user–-interface relationship.

5.3. Operational challenges

5.3.1. Ethical issues for widespread implementation of AI

Understanding and respecting ethical principles is a foundational aspect of clinical practice [92]. The Belmont Report (1979) defines the pillars of biomedical ethics, highlighting key standards that should be adopted for human research [93]. Elements of this need to be taken into consideration throughout the whole AI pipeline process from project design and validation to implementation of the deep learning system. Principles include autonomy, non-maleficence, and equity/justice. Autonomy is based on the understanding that individuals are entitled to make their own healthcare decisions. Non-maleficence – 'do no harm' – ensures that there is no detrimental impact toward patients and their overall clinical outcomes. Equity focuses on fairness with the absence of bias, on a population level [94]. In the context of AI, this ensures that yjr AI system does not accentuate existing health disparities and contribute to inappropriate bias across gender, income, ethnicity, and race. Inappropriate bias can have an expansive impact across the AI pipeline. Representative training datasets should be used during both training and validation in order to ensure inclusivity and minimize pre-existing health bias. The principle of equity is also applicable when discussing accessibility to the deep learning models. Despite a model having high sensitivity, its positive impact is limited to those individuals who have access to the technology, predominately available at large research and healthcare institutes. This in itself further exacerbates inequality and bias based on wealth, location, and accessibility to technology and software [94]. Unintentional bias can also be derived from the value that certain stakeholders seek to gain through AI implementation, which is another benefit of engaging a breadth of stakeholders across the implementation process.

Ethical consideration surrounding the adopting clinician is also important. Dependence on automated machine learning algorithms poses the risk of introducing automation bias along with reducing doctors' skills and confidence in diagnosis and management. Furthermore, concerns regarding the breakdown of the doctor–patient relationship threaten successful implementation. However, it could be argued that AI will maximize clinicians' time to build a rapport with their patients, thus ensuring that their best interests remain at the center of all decision-making [95].

5.3.2. Medicolegal issues & governance

It is crucial that implementation of AI into clinical workflows is done in concordance with appropriate ethical and legal standards. Issues that need addressing include safety, liability, data protection, and privacy concerns [96]. When implementing AI-based systems into ophthalmology workflows, it is important to optimize a pathway for accountability in cases of medical malpractice. For example, if a clinician follows an AI-assisted system in diagnostics and management, but this results in harm to the patient, the question of liability and litigation must be addressed. This needs to be mapped out clearly during the implementation process with uniformity across different hospitals and trusts [96]. The general consensus is that responsibility should still remain with the overseeing clinician, as CDSs are used as a tool to provide additional support and guidance to clinicians. In order to accept that liability, however, clinicians are likely to demand means to confidently evaluate tools and clear guidance to demonstrate their adherence to approved terms of use to demonstrate responsible action on their part. Further complexities and challenges in accountability arise during the implementation of autonomous AI systems. Studies have emphasized the role and responsibility of the manufacturer in determining the quality and safety of the product prior to implementation and hence the need for them to share accountability.

Introduction of AI liability insurance is a potential way of overcoming these issues. This will also help to drive the advances in digital health and encourage integration of high-quality AI by providing companies with ongoing security and support [97]. Other potential solutions to the accountability concern are through the creation of built-in accountability mechanisms into the machine learning algorithm itself, enabling auditing tools with limited explainability [47]. Finally, having consistent standards and an overseeing regulatory body can also help to promote transparency and accountability [98].

Privacy and data protection issues are not unique to AI-enabled tools, and relevant policy and regulatory frameworks are well-established. Patient health data warrant robust adherence to such laws due to the nature of the data. In Europe, The General Data Protection Regulation (GDPR) is responsible for the data protection laws that are enforced. This applies to the rights of processing 'personal data' – which are defined as 'information related to an identified or identifiable person' [96]. Although the conceptual framework of the GDPR is not explicitly designed for AI implementation, there are many components that are relevant and currently used in practice. Particularly within the medical context, 'data concerning health' are distinctly defined as relatiedto the 'physical and mental health of a natural person.' Under the GDPR regulation,

health data, genetic data, and biometrics are all included in a special category of sensitive information. Processing of such data is only permissible provided that there is explicit patient consent or where it is required for public or scientific interest and research [99].

AI governance can be established on a national, regional, organizational, or departmental level depending on the relevant governance structures and feasibility. With the emergence of multiple independent research teams and the expansion of commercially available platforms and algorithms, it is essential that regulation is applied consistently. Collaboration between AI governance bodies, research institutions, and other industry companies is required in order to deploy successful implementation. The formal process involves assessing clinical safety and complexity of implementation and the financial and business assessment of the overall outcome. There is also integration of a feedback process on the selected algorithm, with importance given to the overall clinician experience and accessibility [49]. Governing bodies involved in the process should aim to include individuals from a variety of clinical and non-clinical disciplines including clinical champions, electronic health record managers, data scientists, patients, and legal representatives. This provides collaboration and inclusivity ultimately enabling robust implementation.

6. Conclusion

Artificial intelligence has the potential to revolutionize medical diagnostics, predictive analysis, and management of clinical conditions. Ophthalmology is setting precedents for integration of deep learning into clinical workflows through the value it can add via the image analysis tasks that are so prevalent in ophthalmic clinical practice. Successful translation necessitates a collaborative approach from multiple stakeholders around a structured governance framework, with emphasis on standardization across healthcare providers and equipment and software developers. Safeguards should be put in place to prevent the exacerbation of existing systemic and population biases in healthcare provision. As clinicians have a professional duty toward their patients, concordance with ethical and legal regulations is vital throughout the AI implementation process. We acknowledge that integration of sustainable AI into clinical workflows is a challenging endeavor; however, ophthalmology has the potential to embrace this digital revolution, empowering patients and supporting clinicians in optimizing decision-making.

7. Expert opinion

Artificial intelligence, in particular, machine learning-enabled tools, promises to revolutionize healthcare and diagnostics. As highlighted in the review, successful implementation could have tremendous real-world implications within ophthalmology with regard to imaging diagnosis, treatment guidelines, screening, and predictive analytics. Integration of machine learning-enabled tools into clinical workflow is a challenging process and bridging the current 'AI Chasm' will require a collaborative effort to address issues of transparency, clinical site adaptability, and accountability.

As outcomes of integration will have a direct impact on patient care and well-being, concern regarding clinical adoption is common. These concerns range from systemic technical limitations such as lack of sustainable infrastructure within hospital environments to the personal inherent fear of losing autonomy and clinical expertise. Effective approaches to integration must anticipate and address these issues. Given that a large proportion of health data is embedded within electronic health records, creation of a uniform integration strategy will create a level-playing field. Providing end users with sufficient information about the AI machine learning tool, and knowledge on the data used to generate decisions, will be essential. Other important considerations surround adequate safety and monitoring regulations, which can be applied across different medical hospitals and practices. Without a general agreement on these protective measures, enforcing sustainable widespread changes will be difficult.

Future research should seek to define these barriers further along with effective mitigations to support wide and effective implementation of machine learning-enabled tools into routine patient care. Among other insights, this research will need to support effective collaborations between key stakeholders, explore regulatory needs further, and establish how to ensure these tools are safe. There is also opportunity to improve education processes to enable clinicians, providers, and regulatory bodies to embrace the use of machine learning-enabled tools.

Other promising directions of research include local retraining of machine learning models, with data obtained from local clinical sites. This utilizes the inherent flexibility of these models with the aim of augmenting accuracy and preventing the perpetuation of health and social inequalities [45]. This can help to improve clinician and systemic adoption, ultimately supporting the integration process. Furthermore, availability of "open-source" AI, with data protection policies implemented can also contribute to the scalability of smaller local models [45,100].

There are many lessons that can be learned from the implementation of new technologies outside healthcare. Although much of the focus in recent years has been the personal computer and internet revolutions of the last few decades, perhaps the most instructive example for AI-enabled healthcare may be from the invention and implementation of electric light in the 19th Century. The 'dawning of the electrical age' is commonly said to have occurred in 1882, when Thomas Edison was able to demonstrate electrical lighting for the entire Pearl Street District of Lower Manhattan in New York. The electric lightbulb had, in fact, been invented by many other people and in many different forms in the preceding 25 years [101]. However, the genius of Edison was to develop a network of innovations. Aside from the electric lightbulb itself, these included generators to serve as a source of electricity, a distribution system ('the grid'), a connection system to allow linkage with individual lightbulbs, and a monitoring system ('electricity meters') to allow measurement of consumption. It seems likely that AI-enabled healthcare will likely require similar networks of innovation to achieve its transformative potential in the coming decade.

In summary, this review has provided an overview on the factors that will influence the integration of machine learning-enabled tools into routine ophthalmology clinical workflows. Anticipating and managing these factors will support successful integration with the ultimate aim of streamlining the healthcare eco-system and improving clinical outcomes for individuals and populations.

Funding

Declaration of interest

PA Keane has acted as a consultant for Google, DeepMind, Roche, Novartis, Apellis, and BitFount and is an equity owner in Big Picture Medical. He has received speaker fees from Heidelberg Engineering, Topcon, Allergan, and Bayer. K Balaskas has received speaker fees from Novartis, Bayer, Alimera, Allergan, Roche, and Heidelberg not related to this work; meeting or travel fees from Novartis and Bayer not related to this work; compensation for being on an advisory board from Novartis and Bayer not related to this work; consulting feeds from Novartis and Roche not related to this work; and research support from Apellis, Novartis and Bayer not related to this work. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewer disclosures

Peer reviewers of this manuscript have no relevant financial or other relationships to disclose.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Ophthalmology busiest specialty for third year running [Internet]. Optometry Today [cited 2022 Jun 14]. Available from: https://www. aop.org.uk/ot/professional-support/health-services/2020/10/19/ ophthalmology-busiest-specialty-for-third-year-running.
2. New RCOphth workforce census illustrates the severe shortage of eye doctors in the UK [Internet]. The Royal College of Ophthalmologists [cited 2022 Aug 25]. Available from: https:// www.rcophth.ac.uk/news-views/new-rcophth-workforce-census-illustrates-the-severe-shortage-of-eye-doctors-in-the-uk/.
3. Ting DSJ, Deshmukh R, Said DG, et al. The impact of COVID-19 pandemic on ophthalmology services: are we ready for the aftermath? Ther Adv Ophthalmol. 2020;12:251584142096409.
4. Amisha A, Malik P, Pathania M, et al. Overview of artificial intelligence in medicine. J Fam Med Prim Care. 2019;8(7):2328–2331.
5. The OECD Artificial Intelligence [AI] Principles [Internet]. OECD.AI. [cited 2022 Sep 24]. Available from: https://oecd.ai/en/ai-principles
   • This provides a concise summary of key definitions within the field of AI. This platform (OECD) provides a platform to discuss global policies and governance architecture.
6. Kawamoto K, Houlihan CA, Balas EA, et al. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330 (7494):765.
7. Grzybowski A. Artificial intelligence in ophthalmology. Switzerland AG: Springer; 2021.
8. Tong Y, Lu W, Yu Y, et al. Application of machine learning in ophthalmic imaging modalities. Eye Vis. 2020;7(1):22.

9. Lu W, Tong Y, Yu Y, et al. Applications of artificial intelligence in ophthalmology: general overview. J Ophthalmol. 2018;2018:5278196.
10. Sendak M, D'Archy J, Kashyap S, et al. A path for translation of machine learning products into healthcare delivery. EMJ Innov. 2020;10:172.
    •• This review highlights the complexity of the implementation process – and where integration plays a role within this
11. Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare. Nat Biomed Eng. 2022;6(12):1346–1352.
    •• This provides a comprehensive overview of the concept of self-supervised learning from leading experts in the field. It discusses opportunities for use in healthcare.
12. Lakhani P, Sundaram B. Deep Learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology. 2017;284(2):574–582.
13. Chen L, Bentley P, Rueckert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. NeuroImage Clin. 2017;15:633–643.
14. Menze BH, Jakab A, Bauer S, et al. The multimodal Brain Tumor Image Segmentation Benchmark [BRATS]. IEEE Trans Med Imaging. 2015;34(10):1993–2024.
15. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv: 1711.05225. 2017.
    • This paper provides a leading example of the use of deep learning in a respiratory focused use case.
16. Rajpurkar P, Hannun AY, Haghpanahi M, et al. Cardiologist-level arrhythmia detection with convolutional neural networks arXiv:1707.01836. 2017.
17. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542 (7639):115–118.
    • This study provides evidence of potential use of deep neural networks within Dermatology, by being able to classify skin lesions. This is one of the revolutionary studies within that field.
18. Hammoudi K, Benhabiles H, Melkemi M, et al. Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of COVID-19. J Med Syst. 2021;45(7):75
19. Ghosh P, Azam S, Quadir R, et al. SkinNet-16: a deep learning approach to identify benign and malignant skin lesions. Front Oncol August. 2022;12:931141.
20. 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(6):e271–97.
    •• This is a comprehensive systematic review discussing numerous studies that have compared performance of deep learning models to healthcare professionals.
21. Pontikos N, Basheer K, Kortuem KU, et al. Ten years of optical coherence tomography in ophthalmology: current and future use. Invest Ophthalmol. 2018;59(9):3216–3216.
22. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103 (2):167–175.
23. Jill Hopkins J, Keane PA, Balaskas K. Delivering personalized medicine in retinal care: from artificial intelligence algorithms to clinical application. Curr Opin Ophthalmol. 2020;31(5):329–336.
    •• This paper highlights the potential role of AI in delivery of personalised healthcare, with a focus on ophthalmology
24. ElTanboly A, Ismail M, Shalaby A, et al. A computer-aided diagnostic system for detecting diabetic retinopathy in optical coherence tomography images. Med Phys. 2017;44(3):914–923.
25. 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(9):1342–1350.
26. Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol. 2017;135(11):1170–1176.

•• **This study highlights the use of CNNs in grading Age-Related Macular Degeneration.**

27. 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(22):2402.

28. Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. Nat Med. 2020;26(6):892–899.

• **This is one of the first studies to highlight the role of deep learning in predicting conversion of disease (from milder – to severe form of age-related macular degeneration). This helps to highlight the numerous roles and potential of AI in ophthalmology.**

29. Wang M, Shen LQ, Pasquale LR, et al. An artificial intelligence approach to detect visual field progression in glaucoma based on spatial pattern analysis. Invest Ophthalmol Vis Sci. 2019;60(1):365–375.

30. 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(2):0024.

31. Zhang J, Whebell S, Gallifant J, et al. An interactive dashboard to track themes, development maturity, and global equity in clinical artificial intelligence research. Lancet Digit Health. 2022;4(4):e212–3.

32. AI for human health - dashboard [Internet]. [cited 2022 Nov 28]. Available from: https://aiforhealth.app/

33. Eun J, Han D, Liu X, et al. Teleophthalmology- enabled-and-artificial- intelligence– ready-referral-pathway-for- community-optometry-referrals-of-retinal- disease-[HERMES]:-a-cluster-randomised-superiority-trial-with-a-linked-diagnostic- accuracy-study-HERMES-study-report- 1-study-protocol. BMJ Open. 2022;12:55845.

•• **This is one of the first large scale implementation studies currently taking place.**

34. Beede E, Baylor E, Hersch F, et al. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy [Internet]. Proceedings of the CHI Conference on Human Factors in Computing Systems; 2020 April 25-30; Honolulu HI USA p. 1–12

35. Held LA, Wewetzer L, Steinhäuser J. Determinants of the implementation of an artificial intelligence-supported device for the screening of diabetic retinopathy in primary care - a qualitative study. Health Informatics J. 2022;28(3):14604582221112816.

36. Yap A, Wilkinson B, Chen E, et al. Patients Perceptions of Artificial Intelligence in Diabetic Eye Screening. Asia-Pacific J Ophthalmol. 2022;11(3):287–293.

37. Robinson EL, Guidoboni G, Verticchio A, et al. Artificial intelligence-integrated approaches in ophthalmology: a qualitative pilot study of provider understanding and adoption of AI. Investigative Ophthalmology & Visual Science. Poster session presented at Association for Research in Vision and Ophthalmology; 2022 May 1- 4, Denver (CO)

38. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. 2021;374:n2061.

39. 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(1):39.

40. Artificial Intelligence: LC, Questions TB. Review of Ophthalmology [internet]. 2021. [cited 2022 Aug 23]. Available from: https://www.reviewofophthalmology.com/article/artificial-intelligence-the-big-questions.

41. Product Announcement: optos AI for DR [Internet]. Optos [cited 2022 Oct 6]. Available from: https://www.optos.com/press-releases/2022-optos-ai-for-dr/

42. RetInSight [Internet]. 2022. [cited 2022 Oct 6]. Available from: https://retinsight.com/

43. Gerendas BS, Sadeghipour A, Michl M, et al. Validation of an automated fluid algorithm on real-world data of neovascular age-related macular degeneration over five years. Retina. 2022;42 (9):1673–1682

44. Intelligence in Ophthalmology [Internet]. RetinAI. [cited 2022 Oct 6]. Available from: https://www.retinai.com/

45. Aristidou A, Jena R, Topol EJ. Bridging the chasm between AI and clinical implementation. Lancet. 2022;399(10325):620.

46. Sanchez A, Grandes G, Cortada J, et al. Feasibility of an implementation strategy for the integration of health promotion in routine primary care: a quantitative process evaluation. BMC Fam Pract. 2017;18:1.

47. Paul Y, Hickok E, Sinha A, et al. Artificial Intelligence in the Healthcare Industry in India [internet]. Cis-India. [cited 2022 Aug 20]. Available from: https://cis-india.org/internet-governance/ai-and-healthcare-report

48. Khan SM, Liu X, Nath S, et al. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. Lancet Digit Health. 2021;3(1):e51–66.

49. Daye D, Wiggins WF, Lungren MP, et al. Implementation of Clinical Artificial Intelligence in Radiology: who Decides and How? Radiology. 2022;305(3):212151.

50. What Is Canary Deployment? [Internet]. Semaphore [cited 2022 Aug 25]. Available from: https://semaphoreci.com/blog/what-is-canary-deployment#:~:text=Insoftwareengineering%2Ccanarydeploymenttherestoftheusers.

51. Liu X, Glocker B, McCradden MM, et al. The medical algorithmic audit. Lancet Digit Health. 2022;4(5):e384–97.

52. What Is MLOps? Machine Learning Operations Explained [Internet]. BMC Software Blogs. [cited 2022 Aug 22]. Available from: https://www.bmc.com/blogs/mlops-machine-learning-ops/

53. Lebovitz S, Lifshitz-Assaf H, Levina N, et al. To Engage or Not to Engage with AI for Critical Judgments: how Professionals Deal with Opacity When Using AI for Medical Diagnosis. Organ Sci. 2022;33 (1):126–148.

54. Yang Q, Steinfeld A, Zimmerman J. Unremarkable AI: fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. Proceedings of the CHI Conference on Human Factors in Computing Systems; 2019 May 4- 9;Glasgow Scotland UK p. 1–11

55. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364–1374.

56. Ibrahim H, Liu X, Rivera SC, et al. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. Trials. 2021;22(1):11.

57. Collins G, Dhiman P, Navarro CLA, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open. 2021;11:7.

58. DECIDE-AI Steering Group: Baptiste Vasey;Clifton DA, Collins GS, et al. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence. Nat Med. 2021;27(2):186–187.

•• **This paper summarises the new guidelines that have been released to help support implementation in clinical settings.**

59. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ. 2022;377:e070904.

60. Sendak MP, Gao M, Brajer N, et al. Presenting machine learning model information to clinical end users with model facts labels. Npj Digit Med. 2020;3(1):41.

61. Ethics and accountability in practice [Internet]. Open Government Partnership [cited 2022 Sep 28]. Available from: https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/

62. Miller T. Explanation in artificial intelligence: insights from the social sciences. Artif Intell. 2019;267:1–38.

63. Explainable AI [Internet]. IBM [cited 2022 Aug 19]. Available from: https://www.ibm.com/uk-en/watson/explainable-ai

64. O'Sullivan. Interpretable vs Explainable Machine Learning [Internet]. Towards Data Science [cited 2022 Aug 19]. Available from: https://towardsdatascience.com/interperable-vs-explainable-machine-learning-1fa525e12f48

65. The AI black box problem [Internet]. ThinkAutomation [cited 2022 Aug 19]. Available from: https://www.thinkautomation.com/bots-and-ai/the-ai-black-box-problem/

66. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195.

67. Amann J, Blasimme A, Vayena E, et al. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20(1):310.

68. Human Computer Interaction, Artificial Intelligence and Intelligent Augmentation. ICIS 2020 Proceedings [Internet]. [cited 2022 Aug 13]. Available from: https://aisel.aisnet.org/icis2020/hci_artintel/

69. Singh RP, Hom GL. Abramoff, et al. Current Challenges and Barriers to Real-World Artificial Intelligence Adoption for Healthcare System, Provider, and the Patient Translational Vision Science & Amp: Technology. 2020;9(2):45.

70. Al-Zubaidy M, Hogg J, Maniatopoulos G, et al. Stakeholder Perspectives on Clinical Decision Support Tools to Inform Clinical Artificial Intelligence Implementation: protocol for a Framework Synthesis for Qualitative Evidence. JMIR Res Protoc. 2022;11(4): e33145.

71. Singer SJ, Kellogg KC, Galper AB, et al. Enhancing the value to users of machine learning-based clinical decision support tools: a framework for iterative, collaborative development and implementation. Health Care Manage Rev. 2022;47(2):E21–31.

72. Guzman AL, Lewis SC. Artificial intelligence and communication: a human–machine communication research agenda. New Med Soc. 2019;22:70–86.

73. Sandhu S, Lin AL, Brajer N, et al. Integrating a machine learning system into clinical workflows: qualitative study. J Med Internet Res. 2020;22(11):e22421.

74. Wang S, Hogg J, Devdatta S, et al. Implementation of a machine learning algorithm to enhance primary care effectiveness and equity for peripheral arterial disease: a qualitative study [Internet]. Poster session presented at: Machine Learning for Healthcare conference; 2022 Aug 5-6, Durham, NC 27701

75. Xu W, Dainoff MJ, Ge L, et al. Transitioning to human interaction with AI systems: new challenges and opportunities for HCI professionals to enable human-centered AI. Int J Human–Computer Interact. 2022;39:1–25.

76. Kellogg KC, Sendak M, Balu S. AI on the Front Lines [Internet]. MIT Sloan Manag Rev. 2022 May 4;63(4). Available from https://sloanreview.mit.edu/article/ai-on-the-front-lines/

77. Eyes on FHIR. FHIR for Ophthalmology [Patient Care WG project] - Patient Care [Internet]. Confluence [cited 2022 Oct 7]. Available from: https://confluence.hl7.org/pages/viewpage.action?pageId=82914199

78. Kras A, Oliver W, Gillies MC, et al. Extending interoperability in ophthalmology through Fast Healthcare Interoperability Resource (FHIR). Invest Ophthalmol Vis Sci. 2021;62:1708.

79. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Intern Med. 2021;181(8):1065–1070.

80. Adlung L, Cohen Y, Mor U, et al. Machine learning in clinical decision making. Med. 2021;2(6):642–665.

81. Korot E, Gonçalves MB, Khan SM, et al. Clinician-driven AI in Ophthalmology: resources Enabling Democratization. Curr Opin Ophthalmol. 2021;32(5):445–451.

82. On-premises vs cloud: advantages and disadvantages [Internet]. Empowersuite [cited 2022 Oct 7]. Available from: https://www.empowersuite.com/en/blog/on-premise-vs-cloud

83. De Carlo TE, Romano A, Waheed NK, et al. A review of optical coherence tomography angiography [OCTA]. Int J Retin Vitr. 2015;1(1):5.

84. Bidgood WD, Horii SC, Prior FW, et al. Understanding and using DICOM, the data interchange standard for biomedical imaging. J Am Med Inform Assoc. 1997;4(3):199–212.

85. Caffery LJ, Rotemberg V, Weber J, et al. The role of DICOM in artificial intelligence for skin disease. Front Med. 2021;7:1163.

86. American academy of ophthalmology leads call for ophthalmic equipment manufacturers to standardize digital imaging [Internet]. American Academy of Ophthalmology [cited 2022 Aug 13]. Available from: https://www.aao.org/newsroom/news-releases/detail/american-academy-of-ophthalmology-leads-call-ophth

87. The Royal College of Ophthalmologists endorses the AAO's call for standardisation of digital imaging [Internet]. The Royal College of Ophthalmologists [cited 2022 Aug 13]. Available from: https://www.rcophth.ac.uk/news-views/standardisation-of-digital-imaging/

88. Brusan A, Durmaz A, Ozturk C. Workflow for ensuring DICOM compatibility during radiography device software development. J Digit Imaging. 2021;34(3):717–730.

89. Carter RE, Anand V, Harmon DM, et al. Model drift: when it can be a sign of success and when it can be an occult problem. Intell Med. 2022;6:100058.

90. Hawkins R, Paterson C, Picardi C, et al. Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS) [Internet]. [cited 2022 Nov 28]. Available from: https://www.york.ac.uk/media/assuring-autonomy/documents/AMLASv1.1.pdf

91. ISO. ISO 14971:2019.Medical devices: application of risk management to medical devices [Internet]. [cited 2022 Nov 28]. Available from: https://www.iso.org/standard/72704.html

92. The duties of a doctor registered with the General Medical Council - ethical guidance [Internet]. GMC [cited 2022 Sep 25]. Available from: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice/duties-of-a-doctor

93. The Belmont Report [Internet]. HHS.gov [cited 2022 Sep 25]. Available from: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

94. Abràmoff MD, Cunningham B, Patel B, et al. Foundational Considerations for Artificial Intelligence Using Ophthalmic Images. Ophthalmology. 2022;129(2):e14–32.

95. Abdullah YI, Schuman JS, Shabsigh R, et al. Ethics of Artificial Intelligence in Medicine and Ophthalmology. Asia-Pacific J Ophthalmol. 2021;10(3):289–298.

96. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. Artificial intelligence in healthcare. 2020:295–336. DOI:10.1016/B978-0-12-818438-7.00012-5.

97. Stern AD, Goldfarb A, Minssen T, et al. AI insurance: how liability insurance can drive the responsible adoption of artificial intelligence in health care. NEJM Catal. 2022;3(4):21.

98. Murphy K, Di Ruggiero E, Upshur R, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):14.

99. Health data and data privacy: challenges for data processors under the GDPR [Internet]. Global Data Hub Taylor Wessing's Global Data Hub [cited 2022 Aug 20]. Available from: https://globaldatahub.taylorwessing.com/article/health-data-and-data-privacy-challenges-for-data-processors-under-the-gdpr#:~:text=TheGDPRmakesclearthat,systems%2Cincludingprocessingbythe

100. Harish KB, Price WN, Aphinyanaphongs AY. Open-source clinical machine learning models: critical appraisal of feasibility, advantages, and challenges. JMIR Form Res. 2022;6(4):e33970.

101. Johnson S. How we got to now: six innovations that made the modern world. 1st ed. New York (NY): Penguin; 2015.