# Artificial Intelligence in medicine

## Potential applications and barriers to deployment

*Urvi Sonawane, Matthieu Komorowski*

## 1. Introduction

The application of Artificial Intelligence (AI) to healthcare has gained tremendous momentum in the last decade, offering the potential to streamline patient clinical encounters and improve patient experience, augment clinical decision making, deliver personalised assessments, and reduce healthcare expenditures (Khanna et al. 2022). However, despite these promises, there remains a vast gap between the large number of 'proof of concept' studies published (AI models with restricted clinical application and limited validation) and the relatively few validated and certified AI tools currently deployed in healthcare settings (Esmaeilzadeh 2020; van de Sande et al. 2021; Gómez-González et al. 2020). The reasons behind this lag are complex, multifaceted and vary across settings and healthcare systems, but broadly include technical, ethical, legal, and human factors (Gerke/Minssen/Cohen 2020).

In this chapter, we will delve deeper into the current and potential applications of AI in medicine, exploring the many ways in which this technology can be utilised to improve patient experience and outcomes and/or healthcare effectiveness. Then, we will examine the major barriers preventing deployment and widespread use of these technologies in healthcare settings.

## 2. Survey of current AI applications in medicine

Clinical encounters can broadly be classified into three categories, these being primary care (usually a patient's first point of contact, e.g., general practice, community pharmacy or dental services), secondary care (planned or elective

care – usually in a hospital, urgent and emergency care or mental health care) and tertiary care (highly specialised treatment), along with community health services (see fig. 1 & 2). Because each of these domains presents challenges, bottlenecks and process inefficiencies, AI applications are being developed on all levels of this 'healthcare ecosystem'.

In primary care, AI solutions have been proposed for a number of applications which can be classified into three categories: 1) clinical decision making and care management, 2) predictive modelling and proactive detection of health conditions and 3) administrative tasks (Mistry 2019).

*Figure 1: Overview of the healthcare ecosystem; original figure by NHS Digital (2022).*

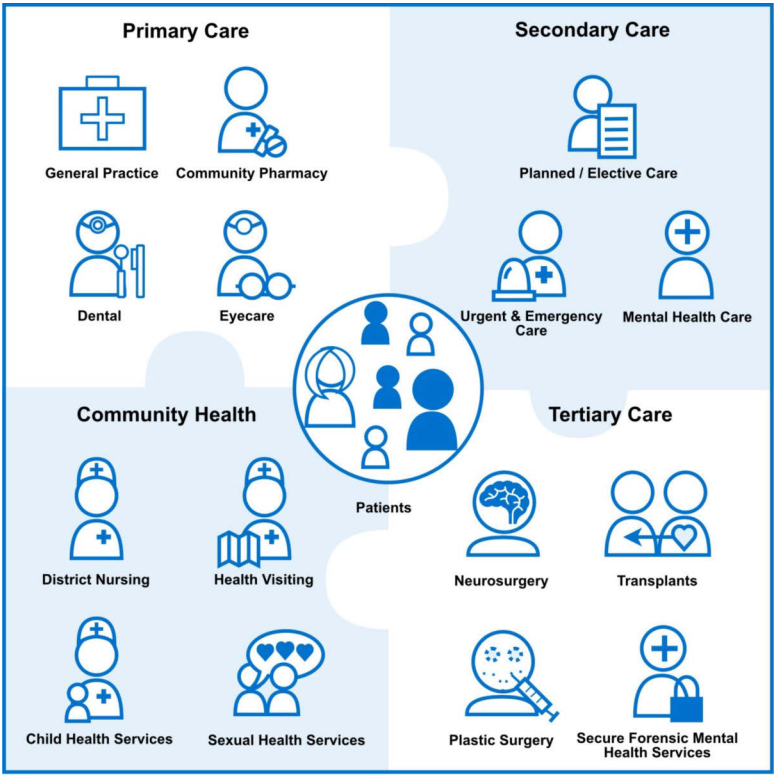
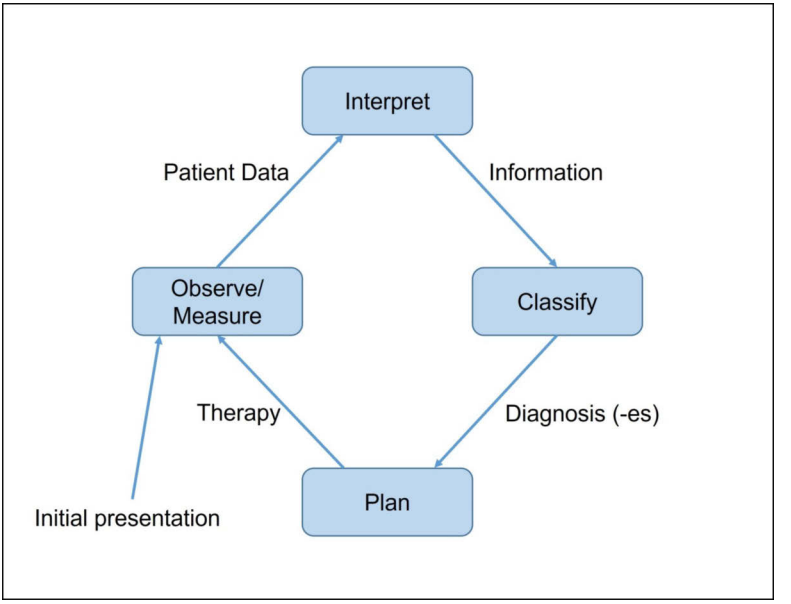


*Figure 2: Summary of a unique clinical encounter. AI applications are being developed at all levels of the healthcare ecosystem and targeting all steps of the clinical encounter; adapted from Groenewegen et al. (2014).*




One example of how clinical decision-making and care management have been influenced by AI in primary care settings is 'Doctor AI', developed by Choi and colleagues (Choi et al. 2016). The predictive model is based on a recurrent neural network and was trained on data from over 260,000 patients over the course of eight years, with the aim to predict diagnoses and medication requirements for the subsequent patient visit (ibid.). With this large dataset, the model achieved 79% recall for diagnosis prediction, which is comparatively higher than other baseline models such as logistic regression or multilayer perceptron (ibid.). In the primary care setting, where integration of specialties and holistic patient care is at the forefront, 'Doctor AI' could contribute to planning subsequent appointment discussion points, thereby assisting primary care clinicians when treating patients that have multiple comorbidities.

Predictive modelling and proactive detection of health conditions has been deployed in diagnosing skin conditions since a seminal publication in *Nature*

in 2017 (Esteva et al. 2017; cf. Liu et al. 2020). Around one in four patients seek out their general practitioner due to skin problems every year (Schofield et al. 2011), and there is an increased demand for a dermatologist review (Eedy 2016). This model used 16,114 anonymised cases from 17 sites to distinguish between 26 skin conditions commonly seen in primary care (Liu et al. 2020). 963 cases were used to validate the system, with the model performing just as well as six board-certified dermatologists and better than six primary care physicians and six nurse practitioners (ibid.). Whilst this is no permanent solution for the increased burden on dermatology in secondary care, it provides an aid for primary care physicians currently facing the impact of strained secondary care.

Administrative tasks have been reported to account for over 50% of general practice time in the UK (Clay/Stern 2015). Furthermore, in the US, one study reported primary care physicians spending nearly two hours on electronic health record tasks for every hour of patient care (Arndt et al. 2017). This indicates an administrative burden on primary care that is not limited to one country. Considering this, Willis and colleagues concluded that there was a potential to 'completely or mostly' automate 44% of administrative tasks carried out by three urban and three rural general practices in England (Willis et al. 2020). This shows massive potential for machine learning to be integrated into primary care. One new development by Microsoft in collaboration with Nuance Communications Inc. is to use conversational AI to provide clinical documentation that writes itself during a clinician-patient encounter (Langston 2019).

In hospital care, AI applications have been developed across the whole patient pathway, from admission prediction, patient triaging, early diagnosis, decision treatment support and outcome prediction. A large research effort is also focusing on auxiliary tasks such as drug discovery, clinical trial enrolment or administrative tasks including appointment scheduling and medical data management.

Progress has been particularly abundant in the field of radiology. As of February 2023, the United States Food and Drug Administration (FDA) has approved 521 machine learning-enabled medical devices, with 71% of them related to radiology (U.S. Food & Drug Administration n.d.). These devices use AI algorithms to analyse images for diagnostics – particularly detecting tumours and identifying patterns in X-rays, CT-scans, MRIs, or tissue samples (Vora et al. 2019). Through the use of AI in medical image analysis, radiologists may potentially provide faster and more accurate diagnoses, which could lead to better patient outcomes. For example, a company developed an AI solution which automatically detects and alerts clinicians of the presence of large vessel

occlusions in the brains of patients suspected to suffer from strokes, with a high sensitivity and specificity, in a real-world prospective setting (Vitellas et al. 2022).

Personalised medicine (also known as precision medicine) is based on the belief that treating, monitoring and preventing diseases must be tailored towards an individual's specific biochemical, physiological, environmental and behavioural profile (Goetz/Schork 2018). The aim is to provide tailored medical care specific to individual patients instead of a broad 'one-size-fits-all' approach, typically provided by expert guidelines (Ruiz-Rodriguez et al. 2022). For example, the management of severe infections in the hospital is dictated by international guidelines such as the "Surviving Sepsis Campaign" (Evans et al. 2021). However, many of the recommendations in such guidelines are based on weak evidence and specific, personalised treatments are not available (Vincent/van der Poll/Marshall 2022). The most likely explanation for this is that sepsis represents a highly heterogeneous patient population, and it is very challenging to identify patients who are more likely to benefit from a specific intervention, for example one targeting components of an immune system (ibid.).

In turn, the concept of data-driven, personalised medicine is becoming increasingly popular, particularly after the COVID-19 pandemic's strain on healthcare provision (Vicente/Ballensiefen/Jönsson 2020). Predictive modelling, which involves using AI algorithms which do not only identify patients at risk for progression of certain diseases, but also predict their responses to treatments, is a particularly promising area of research (Makino et al. 2019; Xu et al. 2021). By accurately predicting a patient's disease progression, healthcare professionals can administer more intensive treatments earlier on, in order to limit long-term disease complications (Makino et al. 2019). This leads to a combination of better patient outcomes coupled with cost cutting, through the reduction of the use of more complex treatments indicated at later stages of disease progression (ibid.). One example for this is taken from a promising study by Makino et al. (2019) where a predictive model was constructed using medical records from over 64,500 diabetic patients to predict diabetic kidney disease progression. The authors suggest that the model can predict diabetic kidney disease progression with 71% accuracy and may reduce the use of haemodialysis, which is known to be a costly intervention in diabetic patients (ibid.; Kent et al. 2015).

AI is also routinely used in natural language processing for popular speech recognition softwares such as 'Google Assistant' and 'Siri' (Google n.d.; Apple

Inc. n.d.). This could also be applied to natural language processing of electronic medical records. Using AI to compile and analyse healthcare records from different staff members could reveal new patterns otherwise almost impossible to spot, let alone diagnose, by human eyes (Mintz/Brodie 2019).

The development of new drugs is imperative for addressing evolving health challenges, such as antibiotic resistance, and AI has the potential to accelerate this process through identifying new drug targets (David et al. 2021). An example of this was shown by Zoffman and colleagues, who used machine learning to search through available antibiotic compounds, eliminate known substances from past projects, and prioritise substances based on factors such as a potency, novelty and availability (Zoffman et al. 2019). This approach can lead to the enhanced discovery of new drugs, particularly in the primary screening stage, as well as the narrowing down and prediction of specific modes of action (ibid.). Drug discovery and development is a complex and expensive process that involves rigorous testing and regulations to ensure safety and efficacy (Chan et al. 2019). One challenge in drug development is ascertaining the toxicology profile of the compound, which can be time-consuming and expensive (Blomme/Will 2016). However, AI systems such as 'DeepTox', which have shown promising accuracy in predicting the toxicology profile of compounds (Mayr et al. 2016), can help to reduce the uncertainty inherent in those processes.

The use of AI in surgical procedures has the potential to significantly improve patient outcome by enhancing precision and accuracy in surgical techniques, with some already being approved by the FDA (Bhandari/Zeffiro/Reddiboina 2020). For instance, AI can be used to identify kidney tumours from bulk CT, allowing surgeons to plan their approach before the procedure commences, or to practice surgical technique in low-risk surgeries (ibid.). However, a systematic review suggested that research surrounding AI in surgery (robot-assisted surgery in particular) is not yet of sufficient quality to safely rely on, primarily due to its limited dataset size (Moglia et al. 2021).

Apart from physical health, there is also growing interest in the application of AI in mental health. Virtual reality and gaming technology, particularly, could help patients with conditions such as depression, bipolar disorder, or chronic pain. By transporting patients into immersive virtual environments, these technologies can provide a safe space for patients to receive psychological therapy and acquire coping mechanisms for their conditions (Hatta et al. 2022; Goudman et al. 2022). While these sessions are currently conducted in the presence of professional staff, the potential for remote sessions should be explored, particularly following the COVID-19 pandemic. This would provide

patients with greater flexibility and accessibility to mental health care, particularly for those living in remote or impoverished areas. Machine learning software also analyses patient responses and feedback during the sessions and learns to adapt more effectively to individual patients, as every manifestation of mental health conditions is unique.

## 3. Limited deployment of AI tools

Despite the extensive list of promising AI applications we detailed above, real world evidence of benefits is lacking for most applications and the validation of AI tools in relevant clinical settings against patient experience and outcomes remains a major challenge.

For example, a rapid search in google scholar for the keywords 'sepsis' and 'prediction' yields over 800,000 results. Comparatively, a 2020 systematic review of the literature focusing on AI identified only 28 published papers, which include mere 3 prospective trials (of which only one was randomised and involved only 142 patients) (Fleuren et al. 2020). Although this number has marginally increased since then with recent publications (e.g., Adams et al. 2022) considering the overall burden of sepsis in the world and the correlated scientific interest generated by sepsis predictions models, the evidence-based benefit of this technology appears worryingly thin.

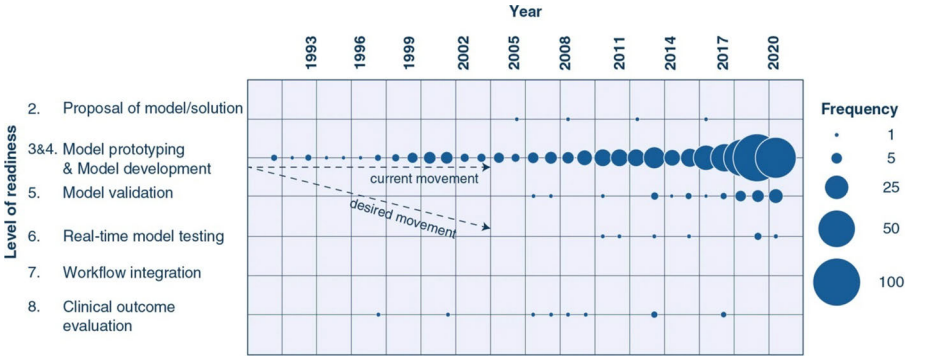
In a 2021 systematic review of AI applications in the intensive care unit, van de Sande and colleagues produced an insightful summary plot (see fig. 3). While there is an increasing number of AI prototypes and early models being developed and trialled, there seems to be a disproportionate disparity when it comes to translating these AI models from production to clinical evaluation. Consequently, the wide gap between the development and clinical implementation of AI tools in intensive care persists, thus limiting the potential benefits that these technologies were intended to achieve (van de Sande et al. 2021).

A group of experts associated with the Joint Research Centre of the European Commission came to a similar conclusion when they reviewed and classified the application of AI in healthcare in terms of current and near-future applications and ethical/social impact (Gómez-González 2020). A novel scale was created to qualify how 'available' healthcare applications were to the public, ranging from 'TAL 0-Unknown status, not considered feasible according to references' to 'TAL 9-Available for the public'. From their systematic search of AI and AI-mediated technologies, most technologies with a positive social impact

were found to have a rating of 'TAL-4-Results of academic/partial projects disclosed', 'TAL-5-Early design of product disclosed' or 'TAL-6-Operational prototype/'first case' disclosed'. This shows that there is still room to drive AI and AI-mediated technologies into the band of TAL-7 to TAL-9, or that there is a fundamental block that needs to be addressed in order to allow more technologies to reach public availability (ibid.).

In the following section, we will explore some of the potential reasons for this phenomenon.

*Figure 3: Number of studies published in Artificial Intelligence in the intensive care unit, according to their level of readiness and year of publication (total number of studies = 494); original figure under CC-BY-NC license by van de Sande et al. (2021).*



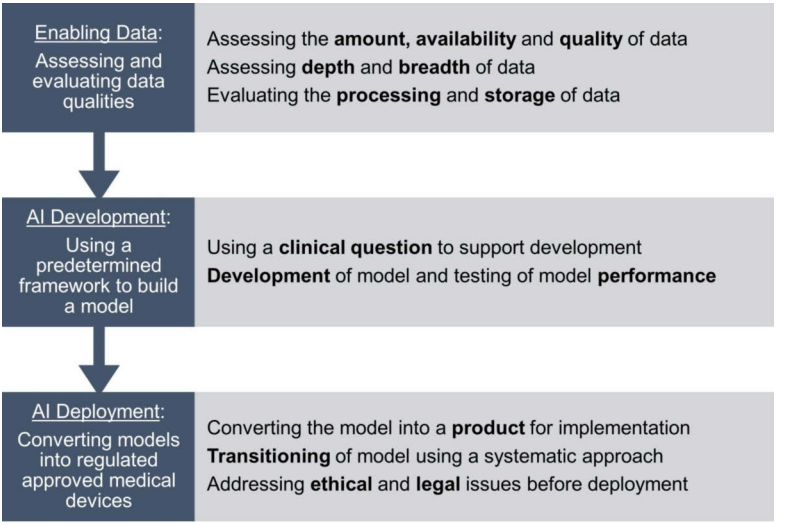

## 4. Challenges to validation and deployment of AI tools in medicine

The process of developing, testing and deploying AI tools in healthcare at scale involves three major steps (see fig. 4: 1) enabling the data, 2) model development and 3) model deployment. Challenges and hurdles are present at each step of this pipeline/process (Mamdani/Slutsky 2021).

The deployment of AI in medicine falls short in comparison to the sheer number of new machine learning inventions proposed in the field of research. This discrepancy can be attributed to several challenges, the first being the lack of model performance in new clinical settings. It is expected for machine learning models to lose some performance due to the differences between real life clinical settings versus the 'developmental environment'. AI models

are often trained with simulated data ortightly controlled parameters, and so models transitioning from the developmental environment to complex real-life clinical settings may face significant differences (Topol 2019). However, this lack of adaptability leads to skepticism about their reliability in crucial clinical judgements that must be accurate. For example, the UK NEWS score was shown to perform poorly in predicting prognosis (AUC 0,6) in a cohort of COVID-19 patients, thus leading to researchers recommending the use of UK NEWS scores as adjuncts to clinical judgment rather than replacements (Colombo et al. 2021).

*Figure 4: Overview of the pipeline for developing, testing, and deploying AI tools in healthcare at scale. This involves three major steps: 1) enabling the data, 2) model development and 3) model deployment; adapted from Mamdani/Slutsky (2021).*




Another hurdle is the scarcity of available data (Ibrahim et al. 2021). AI algorithms require vast amounts of high-quality data to train and test their models. However, obtaining such data is challenging due to patient confidentiality and consent, data sensitivity and lack of cohesive data sharing between the hospitals (Atkin et al. 2021; Kaplan 2016). Electronic health records are often stored in slightly different variations of the same parameter, which creates difficulties in aggregating data and conducting large-scale studies (Holmgren/Adler-

Milstein/McCullough 2018; Dhruva et al. 2020). This can lead to issues with reproducibility and scalability of AI models, as well as difficulties in comparing the performance of different models across different datasets. This limitation means that even the successful algorithms are less suited to be rolled out on a large-scale healthcare service or even across a country (Liang et al. 2022).

Moreover, AI systems are notoriously difficult to integrate within and between systems (this is true within and outside of healthcare) (Baxter/Lee 2021). Currently, most medical AI systems connected to patient data are developed by academic institutions and are not easily usable by external institutions due to the profound discrepancies in IT systems and database structures. Hospitals and small-scale clinics use personalised and/or purpose-built databases to store patient information. Although this makes it easy to navigate through the local area with unique patient demographics, it makes it challenging to adapt the AI code from external institutions to local concept identifiers (Kasparick et al. 2019). Furthermore, hospitals and healthcare systems are often constrained by budget and resource limitations, making it difficult to invest in the necessary infrastructure required to support AI integration (Liang et al. 2022). Therefore, ensuring the AI system is interoperable with other systems, and that data can be shared between different stakeholders in a secure and controlled manner, is challenging.

Many AI models (especially those relying on deep learning) are difficult to interpret and comprehend, which makes it challenging for patients to trust them (Amann et al. 2020). Additionally, patients may not consent to the machine learning software accessing their private data and feed it into ever-changing algorithms due to data security concerns (Atkin et al. 2021). To address this information governance issue, the General Data Protection Regulation (GDPR) in the European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States subjects the obtained AI medical data to strict regulatory and compliance scrutiny. These regulations, which also govern the storage, sharing and use of patient data, can be difficult to navigate through in the context of AI (Liang et al. 2022). All these factors compound into a giant hurdle that has to be overcome.

Another difficulty is the acceptance of this new technology by clinicians. The lack of explainability at this stage makes it challenging to encourage healthcare staff to trust early models. It is possible to observe a gap between expected effect and observed effect even in simple and seemingly innocuous interventions, such as a 'pop-up alert' for acute kidney injury upon the opening of a patient's electronic health record – which shockingly led to a sharp

increase in patient mortality (Wilson et al. 2021). Clinicians' bias with use of technology and evolving (but not confirmed) evidence may be contributing factors. Furthermore, the use of this model for very sick septic patients in the ICU may compound to their lack of trust.

There is no established gold-standard process to demonstrate patient benefit from AI solutions and indeed, there are no recognised best practices for evaluating the efficacy, reliability and safety of commercially available algorithms (Wu et al. 2021). What level of evidence can be accepted by patients, clinicians and regulators? Is retrospective evidence sufficient? Are developers required to conduct multiple randomised trials comparing the standard of care to care supported by their AI solution? Assessment frameworks for the clinical validation of AI have been both proposed (Tsopra et al. 2021; Hawkins et al. 2021; Kickbusch et al. 2020) and surveyed (de Hond et al. 2022), but developing a common set of guidelines for AI model development and implementation remains challenging.

Even those AI applications that have managed to overcome this giant hurdle have issues that need to be considered. The lack of standardisation between AI studies approved for hospital use by regulators (such as the FDA or MHRA) makes it difficult to compare results, mainly due to the varied level of the implementation of studies across different areas of healthcare (Pashkov/Harkusha/Harkusha 2020).

Furthermore, a number of additional human factors must be considered. Healthcare professionals may have a limited or developing understanding and familiarisation of AI tools and would therefore naturally be skeptical of its potential (Gama et al. 2022). This skepticism can make it difficult to integrate AI-based tools into their workflow and practice (Amann et al. 2020). AI tools undoubtedly would also initially add significant expenditure on the already stretched financial healthcare landscape, particularly in the post pandemic period (Kickbusch et al. 2020). One concern of healthcare providers is the legal implications in clinical practice. Healthcare providers may be held liable for potential or actual harm that is caused by AI systems, particularly if they delayed or failed to properly assess or monitor AI's performance. In an era of already burnt-out healthcare staff, the additional responsibility of overseeing the performance of an AI system is unappealing (Gooding/Kariotis 2021). Furthermore, developers of AI systems would also be more cautious when establishing/introducing the software in a position of responsibility given this legal liability (Luxton 2014). The hesitance from both sides is a contributing factor to the lack of implementation of AI software in mainstream healthcare.

## 5. Ethical considerations

The ethical considerations of AI are widely debated, and these concerns are not limited to healthcare. However, certain ethical arguments are particularly pertinent when considering the introduction of AI into mainstream medicine.

AI must provide a real benefit to patients and improve health outcomes and its use must be justified based on patient benefit (Hamet/Tremblay 2017). AI is increasingly being seen as the future of everyday life and financial gain from this cannot be ignored. Deviation from practical patient benefit is certainly possible amidst the desire for investment. Therefore, improving health outcomes should be at the core of AI development in healthcare, which can be done by working in conjunction with patients and healthcare staff.

The presence of discriminative biases in healthcare is undeniable (Ibrahim et al. 2021; Norori et al. 2021). The implication of this, however, could be amplified by AI systems. If they are designed to recognise patterns, these may also perpetuate the existing discrimination in healthcare, leading to further inequality in treatment and health outcomes in patient populations that already experience prejudice and discrimination (Ibrahim et al. 2021; Fletcher/Nakeshima/Olubeko 2021). This ultimately hinders progress towards achieving the desired healthcare equality. For example, an algorithm developed by Gijberts and colleagues using data derived from almost exclusively Caucasian people performed poorly when attempting to predict cardiovascular risk for patients of other ethnicities (e.g., African American and Hispanic ethnicities) (Gijsberts 2015).

AI should be accessible to all patients and should not widen existing health disparities. There is also a potential for AI systems to provide ambiguous or unhelpful answers in critical healthcare situations (Topol 2019). This could again lead to a lack of trust, especially if this happens at the start of implementing the software. It is crucial that the results of this are audited regularly, and the opinions of healthcare staff using the software should be monitored through focus groups and questionnaires to ensure that trust in the software is maintained (Vela et al. 2022).

*Table 1: Summary of the main challenges involved in developing and deploying AI tools in medicine.*

| Step in integrating AI tools in medicine | Challenges involved |
|---|---|
| **Enabling data** | · **Data availability and quality:** effective AI algorithms require large volumes of quality data to refine their models. However, obtaining such data can be difficult, especially when it comes to sensitive medical information (Ibrahim et al. 2021).<br>· **Interoperability and data sharing:** healthcare providers must also ensure that AI systems are interoperable with other systems and that data can be shared between different stakeholders in a standardised, secure and controlled manner. Currently, most systems developed ad-hoc by academic institutions are not usable in external institutions due to profound differences in IT systems and database structures (Baxter/Lee 2021; Kasparick et al. 2019). |
| **Developing AI models** | · **Familiarity with alternate clinical settings:** ML models generally perform sub optimally when deployed in settings dissimilar to those in development. The need for generalisability of AI tools is critical and history is replete with examples of applications that fell short in this regard. For example, the UK NEWS score was shown to perform poorly (AUC 0,6) in a cohort of COVID-19 patients (Colombo et al. 2021).<br>· **Lack of standardisation:** there is a discrepancy in how AI is implemented and used in medicine, which can make it difficult to compare results across different studies and applications (Pashkov/Harkusha/Harkusha 2020; Gama et al. 2022). |

| Step in integrating AI tools in medicine | Challenges involved |
|---|---|
| **Deploying AI models** | · **Bias and discrimination:** AI systems may perpetuate and amplify existing biases in healthcare, leading to unequal treatment and outcomes (Norori et al. 2021).<br>· **Privacy and confidentiality:** data storage, security and protection within AI systems must be compliant with all regulations (e.g., GDPR) (NHS England 2023).<br>· **Responsibility and accountability:** responsibility must be taken for decisions and actions taken by AI systems, particularly in cases where they may cause harm (Gupta/Kamboj/Bag 2021).<br>· **Explainability and transparency:** AI systems may produce results that are difficult to interpret. It's crucial to make sure that the methods and decisions of AI systems are transparent and can be audited (Amann et al. 2020).<br>· **Clinical validity:** evidence based medicine and latest guidelines must be regularly incorporated into AI systems to ensure validity (Crossnohere et al. 2022).<br>· **Clinical utility:** AI must provide a real benefit to patients and improve health outcomes and its use must be justified on the basis of patient benefit (NHS England 2023).<br>· **Equity and access:** AI should be accessible to all patients and should not widen existing health disparities (Gómez-González 2020).<br>· **Regulation and compliance:** AI in medicine is subject to strict regulation and compliance requirements, such as GDPR (EU) and HIPAA (USA), that govern the storage and use of patient data (Crossnohere et al. 2022).<br>· **Legal concerns and liability:** healthcare providers may be held liable for harm caused by AI systems, particularly if they failed to properly assess or monitor the AI's performance (Gupta/Kamboj/Bag 2021).<br>· **Limited understanding of AI by healthcare providers:** healthcare professionals may have a limited understanding of AI, which can make it difficult to integrate AI-based tools into their workflow and practice (Amann et al. 2020). |

Privacy and confidentiality are critical factors that are essential to maintain. Concerns over the risk of patient re-identification have profoundly limited the development of large, publicly available datasets for research. For example, we evaluated what data sources had been used in machine learning models for sepsis resuscitation in the ICU and found that nearly two thirds relied on the same dataset (the MIMIC database) (Johnson et al. 2016). We argue that the benefit of open data sharing outweighs the risks. Indeed, a recent analysis of potential reidentifications of patients in publicly available datasets confirmed that the risk was extremely low (Seastedt et al. 2022). The authors argued that

> the cost – measured in terms of access to future medical innovations and clinical software – of slowing ML progress is too great to limit sharing data through large publicly available databases for concerns of imperfect data anonymization (ibid.).

Data security and patient privacy must be preserved at the phase of model deployment and real-time use, which is a legal requirement and a key aspect of regulatory approval (NHS England 2023).

The security and protection of data is expected by the patient population, yet we may not really know if AI models will be successful in this until they are fully implemented in clinical practice. As per GDPR and NHS medical ethics principles, the patients should be explicitly informed about the use of AI in their care and should also have the autonomy for decision making. If they decided to opt-out of its use, a suitable alternative to the role of the AI software in their care should be offered to all (ibid.).

In situations where AI is involved, accountability and responsibility must be established at all times, including crucial life-or-death decision making. There is also a lack of clarity on which areas of decision making are legally accountable to AI and therefore it is important to identify a clear line of responsibility, including shared responsibility for the decisions and actions taken by AI systems (ibid; Gupta/Kamboj/Bag 2021).

The use of AI in medicine must be supported by clinical evidence and validated through rigorous testing to ensure its accuracy and reliability. Each geographical area has different guidelines supported by different bodies of evidence to suit the varied case mix, and AI systems may be produced in areas with different guidelines. Therefore, it is important to constantly update the treat-

ment algorithms so they comply with constantly evolving medical research and clinical evidence (Crossnohere et al. 2022).

The ethical consideration of financial gain also leads to the point of financial disparities between patients. As mentioned in the first section, the need for AI integration towards 'personalised medicine' would go a long way to make significant savings both by avoiding ineffective treatment costs and better prognosis/quality of life (ibid.).

In conclusion, AI in healthcare holds significant potential to revolutionise the way healthcare operates, from administrative tasks, diagnostics, drug development to surgery. However, despite the many avenues of research that have been and will be explored, there is currently a bottleneck when it comes to the deployment and widespread use of these technologies. This is due to a multitude of factors, including data availability and standardisation, privacy and ethical concerns, clinician and patient skepticism, clinical utility and legal regulations. To overcome these challenges, a collaborative and multidisciplinary approach involving regulatory bodies, healthcare professionals, government entities and patient committees is necessary. This collaboration can produce a clear and regulated framework that will allow innovative and life-changing AI projects to be seamlessly integrated into mainstream healthcare practices.

## List of references

Adams, Roy/Henry, Katherine E./Sridharan, Anirudh/Soleimani, Hossein/Zhan, Andong/Rawat, Nishi/Johnson, Lauren/et al. (2022): "Prospective, Multi-Site Study of Patient Outcomes After Implementation of the TREWS Machine Learning-Based Early Warning System for Sepsis." In: Nature Medicine 28/7, pp. 1455–1460.

Amann, Julia/Blasimme, Alessandro/Vayena, Effy/Frey, Dietmar/Madai, Vince I. (2020): "Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective." In: BMC Medical Informatics and Decision Making 20/310 (https://doi.org/10.1186/s12911-020-01332-6).

Apple Inc. (n.d.): "Siri – Apple (UK).", accessed March 17, 2023 (https://www.apple.com/uk/siri/).

Arndt, Brian G./Beasley, John W./Watkinson, Michelle D./Temte, Jonathan L./Tuan, Wen Jan/Sinsky, Christine A./Gilchrist, Valerie J. (2017): "Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event

Log Data and Time-Motion Observations." In: Annals of Family Medicine 15/5, pp. 419–426.

Atkin, Catherine/Crosby, B./Dunn, K./Price, G./Marston E./Crawford C./O'Hara, M./et al. (2021): "Perceptions of Anonymised Data Use and Awareness of the NHS Data Opt-out Amongst Patients, Carers, and Healthcare Staff." In: Research Involvement and Engagement 7/40 (https://doi.org/10.1186/s40900-021-00281-2).

Baxter, Sally L./Lee, Aaron Y. (2021): "Gaps in Standards for Integrating Artificial Intelligence Technologies into Ophthalmic Practice." In: Current Opinion in Ophthalmology 32/5, pp. 431–438.

Bhandari, Mahendra/Zeffiro, Trevor/Reddiboina, Madhu (2020): "Artificial Intelligence and Robotic Surgery: Current Perspective and Future Directions." In: Current Opinion in Urology 30/1, pp. 48–54.

Blomme, Eric A. G./Will Yvonne (2016): "Toxicology Strategies for Drug Discovery: Present and Future." In: Chemical Research in Toxicology 29/4, pp. 473–504.

Chan, H. C. Stephen/Shan, Hanbin/Dahoun, Thamani/Vogel, Horst/Yuan, Shuguang (2019): "Advancing Drug Discovery via Artificial Intelligence." In: Trends in Pharmacological Sciences (Regular ed.) 40/10, p. 801.

Choi, Edward/Bahadori, Mohammad Taha/Schuetz Andy/Stewart, Walter F./Sun Jimeng (2016): "Doctor AI: Predicting Clinical Events via Recurrent Neural Networks." In: JMLR Workshop and Conference Proceedings 56, pp. 301–318.

Clay, Henry/Stern, Rick (2015): "Making Time in General Practice: Freeing GP Capacity by Reducing Bureaucracy and Avoidable Consultations, Managing the Interface with Hospitals, and Exploring New Ways of Working.", October 2015 (https://thehealthcreationalliance.org/wp-content/uploads/2018/11/Making-Time-in-General-Practice-FULL-REPORT-06-10-15.pdf).

Colombo, Christopher J./Colombo, Rhonda E./Maves, Ryan C./Branche, Angela R./Cohen, Stuart H./Elie, Marie-Carmelle/George, Sarah L./et al. (2021): "Performance Analysis of the National Early Warning Score and Modified Early Warning Score in the Adaptive COVID-19 Treatment Trial Cohort." In: Critical Care Explorations 3/7, e0474.

Crossnohere, Norah L./Elsaid, Mohamed/Paskett, Jonathan/Bose-Brill, Seuli/Bridges, John F. P. (2022): "Guidelines for Artificial Intelligence in Medicine: Literature Review and Content Analysis of Frameworks." In: Journal of Medical Internet Research 24/8, e36823.

David, Liliana/Brata, Anca Monica/Mogosan, Christina/Pop, Christina/Czako, Zoltan/Muresan, Lucian/Ismaiel, Abdulrahman/et al. (2021): "Artificial Intelligence and Antibiotic Discovery." In: Antibiotics (Basel) 10/11, p. 1376.

de Hond, Anne A. H./Leeuwenberg, Artuur, M./Hooft, Lotty/Kant, Ilse M. J./Nijman, Steven W. J./van Os, Hendrikus J. A./Aardoom, Jiksa J./et al. (2022): "Guidelines and Quality Criteria for Artificial Intelligence-Based Prediction Models in Healthcare: A Scoping Review." In: NPJ Digital Medicine 5/2 (https://doi.org/10.1038/s41746-021-00549-7).

Dhruva, Sanket S./Ross, Joseph S./Akar, Joseph G./Caldwell, Brittany/Childers, Karla/Chow, Wing/Ciaccio, Laura/et al. (2020): "Aggregating Multiple Real-World Data Sources Using a Patient-Centered Health Data-Sharing Platform." In: NPJ Digital Medicine 3/60.

Eedy, David (2016): "Dermatology: A Specialty in Crisis." In: Clinical Medicine 16/1, p. 61.

Esmaeilzadeh, Pouyan (2020): "Use of AI-Based Tools for Healthcare Purposes: A Survey Study from Consumers' Perspectives." In: BMC Medical Informatics and Decision Making 20/170 (https://doi.org/10.1186/s12911-020-01191-1).

Esteva, Andre/Kuprel, Brett/Novoa, Roberto A./Ko, Justin/Swetter, Susan M./Blau, Helen M./Thrun, Sebastian (2017): "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." In: Nature 542/7639, pp. 115–118.

Evans, Laura/Rhodes, Andrew/Alhazzani, Waleed/Antonelli, Massimo/Coopersmith, Craig M./French, Craig/Machado, Flávia R./et al. (2021): "Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock 2021." In: Intensive Care Medicine 47/11, pp. 1181–1247.

Fletcher, Richard Ribón/Nakeshimana, Audace/Olubeko, Olusubomi (2021): "Addressing Fairness, Bias, and Appropriate Use of Artificial Intelligence and Machine Learning in Global Health." In: Frontiers in Artificial Intelligence 3 (https://doi.org/10.3389/frai.2020.561802).

Fleuren, Lucas M./Klausch, Thomas L. T./Zwager, Charlotte L./Schoonmade, Linda J./Guo, Tingjie/Roggeveen, Luca F./Swart, Eleonora/et al. (2020): "Machine Learning for the Prediction of Sepsis: a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy." In: Intensive Care Medicine 46/3, pp. 383–400.

Gama, Fábio/Tyskbo, Daniel/Nygren, Jens/Barlow, James/Reed, Julie/Svedberg, Petra (2022): "Implementation Frameworks for Artificial

Intelligence Translation into Health Care Practice: Scoping Review." In: Journal of Medical Internet Research 24/1, e32215.

Gerke, Sara/Minssen, Timo/Cohen, Glenn (2020): "Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare." In: Bohr, Adam/Memarzadeh Kaveh (eds.), Artificial Intelligence in Healthcare, London, San Diego, Cambridge and Oxford: Academic Press, pp. 295–336.

Gijsberts, Crystal M./Groenewegen, Karlijn A./Hoefer, Imo E./Eijkemans, Marinus J. C./Asselbergs, Folkert W./Anderson, Todd J./Britton, Annie R./et al. (2015): "Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events." In: PLoS ONE 10/7, e0132321.

Goetz, Laura H./Schork, Nicholas J. (2018): "Personalized Medicine: Motivation, Challenges, and Progress." In: Fertility and Sterility 109/6, pp. 952–963.

Gómez-González, Emilio/Gomez, Emilia/Márquez-Rivas, Javier/Guerrero-Claro, Manuel/Fernández-Lizaranzu, Isabel/Relimpio-López, María Isabel/Dorado, Manuel E./Mayorga-Buiza, María José/Izquierdo-Ayuso, Guillermo/Capitán-Morales, Luis (2020): Artificial Intelligence in Medicine and Healthcare: A Review and Classification of Current and Near-Future Applications and Their Ethical and Social Impact, arXiv Preprint (https://doi.org/10.48550/arXiv.2001.09778).

Gooding, Piers/Kariotis, Timothy (2021): "Ethics and Law in Research on Algorithmic and Data-Driven Technology in Mental Health Care: Scoping Review." In: JMIR Mental Health 8/6, e24668.

Google (n.d.): "Google Assistant, Your Own Personal Google.", accessed March 17, 2023 (https://assistant.google.com/).

Goudman, Lisa/Jansen, Julie/Billot, Maxime/Vets, Nieke/Smedt, Ann De/Roulaud, Manuel/Rigoard, Philippe/Moens, Maarten (2022): "Virtual Reality Applications in Chronic Pain Management: Systematic Review and Meta-analysis." In: JMIR Serious Games 10/2, e34402.

Groenewegen, André/Tofighy, Azita/Ryvlin, Philippe/Steinhoff, Bernhard J./Dedeken, Peter (2014): "Measures for Improving Treatment Outcomes for Patients with Epilepsy – Results from a Large Multinational Patient-Physician Survey." In: Epilepsy & Behavior 34, pp. 58–67.

Gupta, Shivam/Kamboj, Shampy/Bag, Surajit (2021): "Role of Risks in the Development of Responsible Artificial Intelligence in the Digital Healthcare Domain." In: Information Systems Frontiers (https://doi.org/10.1007/s10796-021-10174-0).

Hamet, Pavel/Tremblay, Johanne (2017): "Artificial Intelligence in Medicine." In: Metabolism, Clinical and Experimental 69, pp. S36–S40.

Hatta, Muhammad Hizri/Sidi, Hatta/Sharip, Shalisah/Das, Srijit/Saini, Suriati Mohamed (2022): "The Role of Virtual Reality as a Psychological Intervention for Mental Health Disturbances During the COVID-19 Pandemic: A Narrative Review." In: International Journal of Environmental Research and Public Health 19/4, p. 2390.

Hawkins, Richard/Paterson, Colin/Picardi, Chiara/Jia, Yan/Calinescu, Radu/Habli, Ibrahim (2021): Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS), arXiv Preprint (https://doi.org/10.48550/arxiv.2102.01564).

Holmgren, A. Jay./Adler-Milstein, Julia/McCullough, Jeffrey (2018): "Are All Certified EHRs Created Equal? Assessing the Relationship Between EHR Vendor and Hospital Meaningful Use Performance." In: Journal of the American Medical Informatics Association 25/6, pp. 654–660.

Ibrahim, Hussein/Liu, Xiaoxuan/Zariffa, Nevine/Morris, Andrew D./Denniston, Alastair K. (2021): "Health Data Poverty: An Assailable Barrier to Equitable Digital Health Care." In: The Lancet Digital Health 3/4, e260–e265.

Johnson, Aistair E. W./Pollard, Tom J./Shen, Lu/Lehman, Li-wei H./Feng, Mengling/Ghassemi, Mohammad/Moody, Benjamin/Szolovits, Peter/Celi, Leo Anthony/Mark, Roger G. (2016): "MIMIC-III, a Freely Accessible Critical Care Database." In: Scientific Data 3/160035 (https://doi.org/10.1038/sdata.2016.35).

Kaplan, Bonnie (2016): "How Should Health Data Be Used?" In: Cambridge Quarterly of Healthcare Ethics 25/2, pp. 312–329.

Kasparick, Martin/Andersen, Björn/Franke, Stefan/Rockstroh, Max/Golatowski, Frank/Timmermann, Dirk/Ingenerf, Josef/Neumuth, Thomas (2019): "Enabling Artificial Intelligence in High Acuity Medical Environments." In: Minimally Invasive Therapy and Allied Technologies 28/2, pp. 120–126.

Kent, Seamus/Schlackow, Iryna/Lozano-Kühne, Jingky/Reith, Christina/Emberson, Jonathan/Haynes, Richard/Gray, Alastair/et al. (2015): "What Is the Impact of Chronic Kidney Disease Stage and Cardiovascular Disease on the Annual Cost of Hospital Care in Moderate-to-Severe Kidney Disease?" In: BMC Nephrology 16/65 (https://doi.org/10.1186/s12882-015-0054-0).

Khanna, Narendra N./Maindarkar, Mahesh A./Viswanathan, Vijay/Fernandes, Jose Fernandes E./Paul, Sudip/Bhagawati, Mrinalini/Ahluwalia, Puneet/et al. (2022): "Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment." In: Healthcare (Basel) 10/12, p. 2493.

Kickbusch, Ilona/Leung, Gabriel M./Bhutta, Zulfiqar A./Matsoso, Malebona Precious/Ihekweazu, Chikwe/Abbasi, Kamran (2020): "COVID-19: How a Virus Is Turning the World Upside Down." In: BMJ 369, m1336.

Langston, Jennifer (2019): "Microsoft and Nuance Join Forces in Quest to Help Doctors Turn Their Focus Back to Patients.", October 17, 2019 (https://news.microsoft.com/source/features/ai/nuance-exam-room-of-the-future/).

Liang, Likeng/Hu, Jifa/Sun, Gang/Hong, Na/Wu, Ge/He, Yuejun/Li, Yong/et al. (2022): "Artificial Intelligence-Based Pharmacovigilance in the Setting of Limited Resources." Drug Safety 45/5, pp. 511–519.

Liu, Yuan/Jain, Ayush/Eng, Clara/Way, David H./Lee Kang/Bui, Peggy/Kanada, Kimberly/et al. (2020): "A Deep Learning System for Differential Diagnosis of Skin Diseases." In: Nature Medicine 26/6, pp. 900–908.

Luxton, David D. (2014): "Recommendations for the Ethical Use and Design of Artificial Intelligent Care Providers." In: Artificial Intelligence in Medicine 62/1, pp. 1–10.

Makino, Masaki/Yoshimoto, Ryo/Ono, Masaki/Itoko, Toshinari/Katsuki, Takayuki/Koseki, Akira/Kudo, Michiharu/et al. (2019): "Artificial Intelligence Predicts the Progression of Diabetic Kidney Disease Using Big Data Machine Learning." In: Scientific Reports 9, 11862.

Mamdani, Muhammad/Slutsky, Arthur S. (2021): "Artificial Intelligence in Intensive Care Medicine." Intensive Care Medicine 47/2, pp. 147–149.

Mayr, Andreas/Klambauer, Günter/Unterthiner, Thomas/Hochreiter, Sepp (2016): "Toxicity Prediction Using Deep Learning." In: Frontiers in Environmental Science 3 (https://doi.org/10.3389/fenvs.2015.00080).

Mintz, Yoav/Brodie, Ronit (2019): "Introduction to Artificial Intelligence in Medicine." In: Minimally Invasive Therapy and Allied Technologies 28/2, pp. 73–81.

Mistry, Pritesh (2019): "Artificial Intelligence in Primary Care." In: British Journal of General Practice 69/686, pp. 422–423.

Moglia, Andrea/Georgiou, Konstantinos/Georgiou, Evangelos/Satava, Richard M./Cuschieri, Alfred (2021): "A Systematic Review on Artificial Intelligence in Robot-assisted Surgery." In: International Journal of Surgery 95, p. 106151.

NHS Digital (2022): "The Healthcare Ecosystem.", last edited October 26, 2022 (https://digital.nhs.uk/developer/guides-and-documentation/introduction-to-healthcare-technology/the-healthcare-ecosystem).

NHS England (2023): "Artificial Intelligence- NHS Transformation Directorate.", February 10, 2023 (https://transform.england.nhs.uk/information-governance/guidance/artificial-intelligence).

Norori, Natalia/Hu, Qiyang/Aellen, Florence Marcelle/Faraci, Francesca Dalia/Tzovara, Athina (2021): "Addressing Bias in Big Data and AI for Health Care: A Call for Open Science." In: Patterns (N.Y.) 2/10, 100347.

Pashkov, Vitlaii M./Harkusha, Andrii O./Harkusha, Yevheniia O. (2020): "Artificial Intelligence in Medical Practice: Regulative Issues and Perspectives." In: Wiadomości Lekarskie 73/12, pp. 2722–2727.

Ruiz-Rodriguez, Juan Carlos/Plata-Menchaca, Erika P./Chiscano-Camón, Luis/Ruiz-Sanmartin, Adolfo/Pérez-Carrasco, Marcos/Palmada, Clara/Ribas, Vicent/Martínez-Gallo, Mónica/et al. (2022): "Precision Medicine in Sepsis and Septic Shock: From Omics to Clinical Tools." In: World Journal of Critical Care Medicine 11/1, pp. 1–21.

Schofield, Julia K./Fleming, D./Grindlay, D./Williams, H. (2011): "Skin Conditions are the Commonest New Reason People Present to General Practitioners in England and Wales." In: British Journal of Dermatology 165/5, pp. 1044–1050.

Seastedt, Kenneth P./Schwab, Patrick/O'Brien, Zach/Wakida, Edith/Herrera, Karen/Marcelo, Portia Grace F./Agha-Mir-Salim, Louis/Frigola, Xavier Borrat/Ndulue, Emily Boardman/Marcelo, Alvin/et al. (2022): "Global Healthcare Fairness: We Should Be Sharing More, Not Less, Data." In: PLOS Digital Health 1/10, e0000102.

Topol, Eric J. (2019): "High-Performance Medicine: The Convergence of Human and Artificial Intelligence." In: Nature Medicine 25/1, pp. 44–56.

Tsopra, Rosy/Fernandez, Xose/Luchinat, Claudio/Alberghina, Lilia/Lehrach, Hans/Vanoni, Marco/Dreher, Felix/et al. (2021): "A Framework for Validating AI in Precision Medicine: Considerations from the European ITFoC Consortium." In: BMC Medical Informatics and Decision Making 21/274 (https://doi.org/10.1186/s12911-021-01634-3).

U.S. Food & Drug Administration (n.d.): "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices.", accessed March 17, 2023 (https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices).

van de Sande, Davy/van Genderen, Michel E./Huiskens, Joost/Gommers, Diederick/van Bommel, Jasper (2021): "Moving From Bytes to Bedside: A Systematic Review on the Use of Artificial Intelligence in the Intensive Care Unit." In: Intensive Care Medicine 47/7, pp. 750–760.

Vela, Monica B./Erondu, Amarachi I./Smith, Nichole A./Peek, Monica E./Woodruff, James N./Chin, Marshall H. (2022): "Eliminating Explicit and Implicit Biases in Health Care: Evidence and Research Needs." In: Annual Review of Public Health 43/1, pp. 477–501.

Vicente, Astrid M./Ballensiefen, Wolfgang/Jönsson, Jan-Ingvar (2020): "How Personalised Medicine Will Transform Healthcare by 2030: The ICPerMed Vision." In: Journal of Translational Medicine 18/1, p. 180.

Vincent, Jean-Louis/van der Poll, Tom/Marshall, John C. (2022): "The End of 'One Size Fits All' Sepsis Therapies: Toward an Individualized Approach." In: Biomedicines 10/9 p. 2260.

Vitellas, Carol A./Mannix, Nicholas C./Nimjee, Shahid M./Shujaat, Mohammad T./Heaton, Sharon/Lee, Vivien H. (2022): "Abstract 130: Real World Experience With Viz.AI Automated Large Vessel Occlusion Detection." In: Stroke 53/Suppl_1, A130.

Vora, Amisha/Malik, Paras/Pathania, Monika/Rathaur, Vyas Kumar (2019): "Overview of Artificial Intelligence in Medicine." In: Journal of Family Medicine and Primary Care 8/7, pp. 2328–2331.

Willis, Matthew/Duckworth, Paul/Coulter, Angela/Meyer, Eric T./Osborne, Michael (2020): "Qualitative and Quantitative Approach to Assess the Potential for Automating Administrative Tasks in General Practice." In: BMJ open 10, e032412.

Wilson, F. Perry/Martin, Melissa/Yamamoto, Yu/Partridge, Caitlin/Moreira, Erica/Arora, Tanima/Biswas, Aditya/Feldman, Harold/Garg, Amit X./Greenberg, Jason H. (2021): "Electronic Health Record Alerts for Acute Kidney Injury: Multicenter, Randomized Clinical Trial." In: BMJ 372, m4786.

Wu, Eric/Wu, Kevin/Daneshjou, Roxana/Ouyang, David/Ho, Daniel E./Zou, James (2021): "How Medical AI Devices Are Evaluated: Limitations and Recommendations from an Analysis of FDA Approvals." In: Nature Medicine 27/4, pp. 582–584.

Xu, Ying/Su, Guan-Hua/Ma, Ding/Xiao, Yi/Shao, Zhi-Ming/Jiang, Yi-Zhou (2021): "Technological Advances in Cancer Immunity: From Immunogenomics to Single-Cell Analysis and Artificial Intelligence." In: Signal Transduction and Targeted Therapy 6/312 (https://doi.org/10.1038/s41392-021-00729-7).

Zoffman, Sannah/Vercruysse, Maarten/Benmansour, Fethallah/Maunz, Andreas/Wolf, Luise/Marti, Rita Blum/Heckel, Tobias/et al. (2019): "Machine Learning-Powered Antibiotics Phenotypic Drug Discovery." In: Scientific Reports 9, 5013 (https://doi.org/10.1038/s41598-019-39387-9).