**The Applications of Artificial Intelligence in Medical Diagnosis**

**Introduction**

With the prevalence and usage of AI across a multitude of industries, there is vast potential for its application in healthcare, particularly regarding medical diagnostics. This is not to say that there has not been a long history of AI in healthcare. Although slow to adopt AI initially (as stated by Kaul et al., 2020), the development of AI in medical diagnosis began as early as the 1970’s. For example, in 1978, Weiss et al. (1978) developed a glaucoma consultation computer.

More recently, there have been a plethora of research studies suggesting that AI can not only perform medical diagnoses but also outperform humans in both accuracy and expediency (Davenport & Kalakota, 2019). An example application of AI could be analysing medical images such as x-rays and ultrasounds to identify and diagnose diseases in radiology (Hosny et al., 2018). An additional use case could be in analysing substantial amounts of data, including patient data, medical history, and demographic information (Kumar et al., 2022), to define and diagnose various conditions.

However, when analysing data and training AI models, there is a danger of introducing bias. For example, in a study that analysed blood tests that diagnosed liver disease, there was a notable difference in accuracy between men and women, and it was found that the results were due to inequalities in care (Straw & Wu, 2022).

Across the literature discussed thus far, there are three clear themes:

* History and development of AI in medical diagnostics
* The applications and efficacy of AI in medical diagnostics
* The potential for bias in AI in medical diagnostics

This review will focus on these themes that have emerged throughout the reviewed literature, and while there is scope for focusing on any one of these themes in more depth, this review aims to provide a broader overview of the common concepts within the AI in Medical Diagnosis topic. Additionally, there may be other themes that fall within the overarching topic but will remain outside of this review, as the above appear to be the most prevalent.

**History and development of AI in medical diagnostics**

The beginnings of the use of AI in medicine begin with the first steps towards artificial neural networks (ANN) in 1943 when Walter Pitts developed the first models of neural networks (McCulloch & Pitts, 1943). However, it was not until the 1970s that there were any significant innovations in relation to AI in medicine (Kulikowski, 2015).

In these initial stages, from the available literature at the time, most AI models were limited to programmes that advised in consultations (Shortliffe et al., 1973; Weiss et al., 1978; Pople, 1975; and Pauker et al., 1976).

This interest in AI applications for biomedical purposes in this period culminated in 1975 at the First Rutgers AIM Workshop (Kulikowski, 2015), where many of the above models were presented and discussed. This was later followed by the development of MYCIN (van Melle, 1978) and the use of PIP (Present Illness Program) to assist in the diagnosis and evaluation of bacterial pathogens and oedema, respectively.

This trend of AI systems to assist in diagnosis via human input but development stagnated in the ‘AI winter’ from 1974–1993 (Tang, 2020), although Kaut et al. (2020) extend this as far as two thousand. For example, it was not until around 2011 when AI was used in diagnostic imaging in radiation oncology (Magalhães Barros Netto et al., 2012) and has been developed to the point that radiology is less about human perception and more objective (Chartrand et al., 2017). Although this might indicate that human clinicians may become surplus to the requirements for diagnosis in the future, Sezgin (2023) posits that AI is merely there to complement doctors rather than replace them.

**Applications and efficacy of AI in medical diagnostics**

There are many applications of AI in medicine as demonstrated by a systematic literature review conducted by Kumar et al. (2022) (see *figure 1* below).

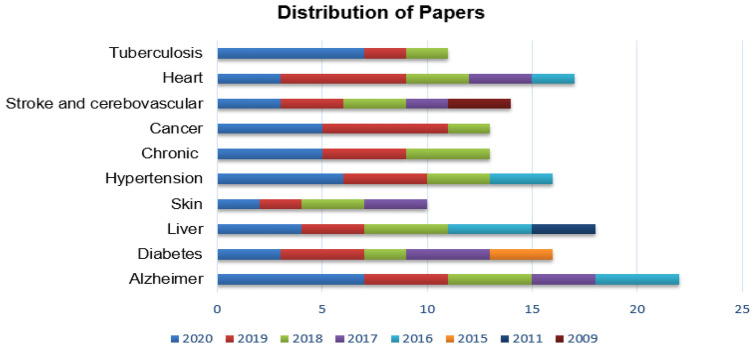


Figure 1 - Distribution of published papers using AI in disease diagnosis (Kumar et al., 2022)

This portion of the review will look at specific applications and their efficacy rather than a more holistic view, which can be observed in the systematic literature review.

One such application is the use of deep learning to detect COVID-19 using chest X-rays (Brunese et al., 2020). Because patients infected with COVID-19 may be asymptomatic (Romagnani et al., 2020), the swab test could prove ineffective as the test is only carried out by those showing symptoms (Brunese et al., 2020). Thus, the aim is to provide a more robust method of diagnosis.

In this study, they used a CNN model using VGG-16 architecture to pass through a collection of 6,523 chest X-ray images from three different datasets: one with COVID-19 patients, one with other pulmonary diseases, and one with healthy patients.

The first model produced aimed to discriminate between healthy patients and pulmonary disease patients (including COVID-19), and the second model aimed to discriminate between COVID-19 X-rays and pulmonary disease X-rays. Both models have an accuracy of 0.96 or above and can detect the presence of disease or COVID-19 (depending on the model) in approximately 2.5 seconds.

Whilst unlikely to have much use due to the impracticalities of having potential COVID-19 patients X-rayed, it is nonetheless impressive regarding the model's accuracy and could have further uses for identifying other pulmonary diseases.

Another specific application is in kidney disease, specifically immunoglobulin A nephropathy (IgAN) and the prediction of patients with the ailment developing end-stage renal disease (ESRD) (Liu et al., 2018). Before this model was developed, this area had not been well studied in Asian patients, as the paper states.

A random forest model was developed with six predictors. This model takes the characteristics of a case and based on the input values, will proceed to a different branch of a decision tree.

Once the initial model had been produced using importance ranking methods, the crucial predictors could be established. The model was then verified with a logistic regression analysis. By knowing the importance of factors, new models could be created with additional key factors to help improve accuracy. The final model was observed to have an AUC of 95.45% and established that factors such as Oxford-MEST scores and C3 staining are important in improving the accuracy of predictions for ESRD.

While limited in scope to the Asian population, the tests and methods may be applicable to other ethnic populations as the predictors used are universal.

**Bias in AI medical diagnostics**

The literature in this area tends to focus on sex, gender, and ethnicity. The topic is not limited to bias that originates from data that causes inaccuracies that discriminate due to the datasets AI is trained on misrepresenting or excluding classes of people (Norori et al., 2021), but also in research practices and workplaces where AI is involved (Buslón et al., 2023) as well as the contrary argument to the former that AI can help alleviate human biases (Brown et al., 2023).

It is pertinent to define bias in the context of AI as “a difference in performance between subgroups for a predictive task” (Yang et al., 2023). This follows the literature regarding such AI reports on underrepresented groups in terms of ethnicity and gender.

Indeed, much of the literature tends to focus on its existence or proving that it exists. For example, AlHasan (2021) collates several studies pertaining to racial, gender, socioeconomic, and linguistic bias resulting from the datasets the AI is trained on being prejudiced and consequently skews the AI model to show those prejudices in results. One such example is AI in plastic surgery, where attractiveness determined by western countries conflicts with perceived beauty in other countries (Koimizu et al., 2019).

Further, there are issues with underdiagnosis for females, black patients, and patients of lower socioeconomic backgrounds resulting from AI algorithms that incorrectly label patients with a disease as healthy (Seyyed-Kalantari et al., 2021). A specific example is demonstrated in a study that found neural networks that utilised imbalanced X-ray image datasets (in terms of representation of gender) to diagnose thoracic diseases, resulting in underdiagnosis (Larrazabal et al., 2020). Additionally, there is evidence of racial bias in the US health system, where a widely used AI algorithm found black patients were identified for extra care by more than half compared to white patients due to the algorithm predicting healthcare costs (Obermeyer et al., 2019). This emphasises the importance of understanding the underlying datasets and their potential for bias.

Bias emanating from AI can only serve to exacerbate issues in healthcare equity and lead to further misdiagnosis and incorrect care (Celi et al., 2022). There is some agreement that AI’s initial purpose was to help alleviate concerns about bias originating from humans and improve health equity (Brown et al., 2023; Nazer et al., 2023). Because of the increase in literature on proving the existence of bias in AI, there has been subsequent literature focusing on mitigating bias.

Strategies to combat bias in AI algorithms include comprehensive checklists and tools such as CLAIM (Checklist for AI in Medical Imaging) (Mongan et al., 2020) and PROBAST (Prediction Model Risk of Bias Assessment Tool) (Wolff et al., 2019). These tools alone predominately help authors assess AI algorithms but do not help address bias (Nazer et al., 2023). A framework has been developed for more equitable healthcare in AI; however, this only focuses on racial discrimination and does not cover other underrepresented groups (Dankwa-Mullan et al., 2021). Nazer et al. (2023), considering this, have proposed a more comprehensive checklist at each step of the development of an AI algorithm, from the initial framing of the problem to its implementation. It asks developers questions on what subgroups make up the population, what are the data sources, and were de-biassing techniques adopted? Other literature suggests that, given an expert aware of the bias, keeping humans involved can be an important strategy in mitigating bias, amongst other strategies (Mittermaier et al., 2023).

Additionally, there have been attempts at regulation from organisations such as the US Food and Drug Administration (FDA) (Muehlematter et al., 2023) and the World Health Organization (WHO) (World Health Organization, 2023).

**Strengths, limitations, and discrepancies**

Because of the broad nature of this review, each theme will be discussed in terms of its strengths and weaknesses.

The literature on the history and development of AI in medical diagnosis was comprehensive but not abundant. There is no definitive timeline for early adoption of AI, but there is some literature available from the 1970s onwards. However, these articles are not easily accessible, and only a brief overview could be included in the review. The content coincided with the author's personal interests in technology and history. Despite this, there is enough information provided to give an idea of how AI has been adopted over time.

The biggest strengths of the literature come from comprehensive discussions and analyses of the current and future applications of AI in medical diagnosis. This theme allowed for both breadth and depth. For example, there could be a literature review of AI in liver disease diagnosis due to the abundance of literature available. Due to the sheer volume of articles, I opted for a broader overview to provide an idea of the use cases and leave it to the reader to research further if a particular disease interests them.

A limitation of the literature regarding bias is that there are limited papers that cover both specific bias and general bias. This may be due to the difficulty of accessing datasets that in-use AI has been trained on, as well as not knowing fully how the AI model was developed. Additionally, most literature (including that from organisations such as the FDA and WHO) bringing forth suggestions and guidance on avoiding bias does not need to be followed. If the issue is to be taken seriously, there need to be rules in place to regulate AI in this area, as lives are at risk due to underdiagnosis and subsequent lack of care.

**Conclusion**

The literature for AI in medical diagnosis is vast, and the role of AI in healthcare is constantly expanding (Ahmad et al., 2021). Papers range from the study of the history of AI in medicine to potential or current-specific applications. Because this area is so active, there is an abundance of content to explore and research, and one literature review could not easily contain everything. Thus, this literature review has aimed to achieve a brief overview of the various themes to guide the reader to areas that may be of interest to them.

The history of medical diagnosis starts in the 1900s and is continuing to grow. Starting out with more basic AI that predicted a diagnosis based on a series of inputs to CNN models based on millions of images such as X-rays to identify diseases as accurately (and sometimes more accurately) as clinicians (Goh, 2023).

It is vital to maintain awareness of bias when evaluating models, and as stated above, without some regulation, AI may exacerbate existing biases if not kept in check. However, I do acknowledge that regulation in this regard may limit innovation in AI, but it is necessary for the accuracy and equity of healthcare diagnoses.

**Word Count: 2199**

**References**

Ahmad, Z. *et al.* (2021) ‘Artificial Intelligence (AI) in medicine, current applications, and future role with special emphasis on its potential and promise in pathology: Present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review’, *Diagnostic Pathology*, 16(1). doi:10.1186/s13000-021-01085-4.

AlHasan, A. (2021) ‘Bias in medical artificial intelligence’, *The Bulletin of the Royal College of Surgeons of England*, 103(6), pp. 302–305. doi:10.1308/rcsbull.2021.111.

Brown, C. *et al.* (2023) ‘Breaking bias: The role of Artificial Intelligence in improving clinical decision-making’, *Cureus* [Preprint]. doi:10.7759/cureus.36415.

Buslón, N. *et al.* (2023) ‘Raising awareness of sex and gender bias in artificial intelligence and health’, *Frontiers in Global Women’s Health*, 4. doi:10.3389/fgwh.2023.970312.

Brunese, L. *et al.* (2020) ‘Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays’, *Computer Methods and Programs in Biomedicine*, 196, p. 105608. doi:10.1016/j.cmpb.2020.105608.

Celi, L.A. *et al.* (2022) ‘Sources of bias in artificial intelligence that perpetuate healthcare disparities—a global review’, *PLOS Digital Health*, 1(3). doi:10.1371/journal.pdig.0000022.

Dankwa-Mullan, I. *et al.* (2021) ‘A proposed framework on integrating health equity and racial justice into the Artificial Intelligence Development Lifecycle’, *Journal of Health Care for the Poor and Underserved*, 32(2S), pp. 300–317. doi:10.1353/hpu.2021.0065.

Davenport, T. and Kalakota, R. (2019) ‘The potential for artificial intelligence in Healthcare’, *Future Healthcare Journal*, 6(2), pp. 94–98. doi:10.7861/futurehosp.6-2-94.

Goh, V. (2023) *AI trained on X-rays can diagnose medical issues as accurately as doctors*, *King’s College London*. Available at: https://www.kcl.ac.uk/news/ai-trained-on-x-rays-can-diagnose-medical-issues-as-accurately-as-doctors (Accessed: 02 January 2024).

Hosny, A. *et al.* (2018) ‘Artificial Intelligence in radiology’, *Nature Reviews Cancer*, 18(8), pp. 500–510. doi:10.1038/s41568-018-0016-5.

Kaul, V., Enslin, S. and Gross, S.A. (2020) ‘History of artificial intelligence in medicine’, *Gastrointestinal Endoscopy*, 92(4), pp. 807–812. doi:10.1016/j.gie.2020.06.040.

Koimizu, J., Numajiri, T. and Kato, K. (2019) ‘Machine Learning and Ethics in Plastic surgery’, *Plastic and Reconstructive Surgery - Global Open*, 7(3). doi:10.1097/gox.0000000000002162.

Kononenko, I. (2001) ‘Machine learning for medical diagnosis: History, state of the art and perspective’, *Artificial Intelligence in Medicine*, 23(1), pp. 89–109. doi:10.1016/s0933-3657(01)00077-x.

Kulikowski, C.A. (2015) ‘An opening chapter of the first generation of Artificial Intelligence in medicine: The first Rutgers Aim Workshop, June 1975’, *Yearbook of Medical Informatics*, 24(01), pp. 227–233. doi:10.15265/iy-2015-016.

Kumar, Y. *et al.* (2022) ‘Artificial Intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda’, *Journal of Ambient Intelligence and Humanized Computing*, 14(7), pp. 8459–8486. doi:10.1007/s12652-021-03612-z.

Larrazabal, A.J. *et al.* (2020) ‘Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis’, *Proceedings of the National Academy of Sciences*, 117(23), pp. 12592–12594. doi:10.1073/pnas.1919012117.

Liu, Y. *et al.* (2018) ‘Prediction of ESRD in IGA nephropathy patients from an Asian cohort: A random forest model’, *Kidney and Blood Pressure Research*, 43(6), pp. 1852–1864. doi:10.1159/000495818.

Magalhães Barros Netto, S. *et al.* (2012) ‘Automatic segmentation of lung nodules with growing neural gas and support vector machine’, *Computers in Biology and Medicine*, 42(11), pp. 1110–1121. doi:10.1016/j.compbiomed.2012.09.003.

McCulloch, W.S. and Pitts, W. (1943) ‘A logical calculus of the ideas immanent in nervous activity’, *The Bulletin of Mathematical Biophysics*, 5(4), pp. 115–133. doi:10.1007/bf02478259.

Mittermaier, M., Raza, M.M. and Kvedar, J.C. (2023) ‘Bias in AI-based models for medical applications: Challenges and mitigation strategies’, *npj Digital Medicine*, 6(1). doi:10.1038/s41746-023-00858-z.

Mongan, J., Moy, L. and Kahn, C.E. (2020) ‘Checklist for Artificial Intelligence in Medical Imaging (Claim): A guide for authors and reviewers’, *Radiology: Artificial Intelligence*, 2(2). doi:10.1148/ryai.2020200029.

Muehlematter, U.J., Bluethgen, C. and Vokinger, K.N. (2023) ‘FDA-cleared Artificial Intelligence and machine learning-based medical devices and their 510(k) predicate networks’, *The Lancet Digital Health*, 5(9). doi:10.1016/s2589-7500(23)00126-7.

Nazer, L.H. *et al.* (2023) ‘Bias in artificial intelligence algorithms and recommendations for mitigation’, *PLOS Digital Health*, 2(6). doi:10.1371/journal.pdig.0000278.

Norori, N. *et al.* (2021) ‘Addressing bias in big data and AI for health care: A call for open science’, *Patterns*, 2(10), p. 100347. doi:10.1016/j.patter.2021.100347.

Obermeyer, Z. *et al.* (2019) ‘Dissecting racial bias in an algorithm used to manage the health of populations’, *Science*, 366(6464), pp. 447–453. doi:10.1126/science.aax2342.

Pauker, S.G. *et al.* (1976) ‘Towards the simulation of clinical cognition’, *The American Journal of Medicine*, 60(7), pp. 981–996. doi:10.1016/0002-9343(76)90570-2.

Pople, H.E., Myers, J.D. and Miller, R.A., 1975, September. DIALOG: A model of diagnostic logic for internal medicine. In *IJCAI* (Vol. 4, pp. 848-855).

Romagnani, P. *et al.* (2020) ‘The COVID-19 infection: Lessons from the Italian experience’, *Journal of Public Health Policy*, 41(3), pp. 238–244. doi:10.1057/s41271-020-00229-y.

Seyyed-Kalantari, L. *et al.* (2021) ‘Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations’, *Nature Medicine*, 27(12), pp. 2176–2182. doi:10.1038/s41591-021-01595-0.

Sezgin, E. (2023) ‘Artificial Intelligence in Healthcare: Complementing, not replacing, doctors and healthcare providers’, *DIGITAL HEALTH*, 9. doi:10.1177/20552076231186520.

Shortliffe, E.H. *et al.* (1973) ‘An artificial intelligence program to advise physicians regarding antimicrobial therapy’, *Computers and Biomedical Research*, 6(6), pp. 544–560. doi:10.1016/0010-4809(73)90029-3.

Straw, I. and Wu, H. (2022) ‘Investigating for Bias in healthcare algorithms: A sex-stratified analysis of supervised machine learning models in liver disease prediction’, *BMJ Health Care Inform*, 29(1). doi:10.1136/bmjhci-2021-100457.

Tang, X. (2020) ‘The role of Artificial Intelligence in medical imaging research’, *BJR|Open*, 2(1), p. 20190031. doi:10.1259/bjro.20190031.

van Melle, W. (1978) ‘MYCIN: A knowledge-based consultation program for infectious disease diagnosis’, *International Journal of Man-Machine Studies*, 10(3), pp. 313–322. doi:10.1016/s0020-7373(78)80049-2.

Weiss, S., Kulikowski, C.A. and Safir, A. (1978) ‘Glaucoma consultation by computer’, *Computers in Biology and Medicine*, 8(1), pp. 25–40. doi:10.1016/0010-4825(78)90011-2.

Wolff, R.F. *et al.* (2019) ‘PROBAST: A tool to assess the risk of bias and applicability of Prediction model studies’, *Annals of Internal Medicine*, 170(1), p. 51. doi:10.7326/m18-1376.

World Health Organization. (‎2023)‎. Regulatory considerations on artificial intelligence for health. World Health Organization. <https://iris.who.int/handle/10665/373421>.

Yang, J. *et al.* (2023) ‘An adversarial training framework for mitigating algorithmic biases in clinical machine learning’, *npj Digital Medicine*, 6(1). doi:10.1038/s41746-023-00805-y.