To what extent do the use of artificial intelligence systems in medicine contribute to the socio-economic condition of society?

Table of Contents

Abstract

Implementation of artificial intelligence systems in medicine is very limited but highly effective where utilised. Use of these systems in smart devices such as smart watches to detect atrial fibrillation, a common heart condition, is saving millions of dollars in health care expenditure and saving hundreds of lives. These traits are broadly applicable to all of the roughly 26 medicinal artificial intelligence systems in use, with around half of these being used for image analysis, although the economic benefit and number of lives improved vary depending on the disease each system targets. Saving or improving the quality of lives through effective diagnosis and treatment is enormously economically beneficial as it reduces healthcare expenditure, enables a person to work better and for longer (thus contributing more to the economy) not to mention the fact that it enables doctors to spend more time on more advanced cases. The social benefits are also obvious as saving lives and enhancing the quality of life is massively beneficial to the patient and those close to the patient. As a result of the large number of lives saved or improved, the socio-economic impact for implemented artificial intelligence systems is large per system although small in the wider picture of medicine due to the small number of systems currently in use. Research into artificial intelligence in healthcare also has a positive current socio-economic impact as it has directly led to the systems that exist today, although this impact is far smaller than the impact of implemented systems. Far more systems are being researched and many of them prove that AI can perform at the level of experts which encourages more investment and research into the field which will lead to large long-term socio-economic advances. However, current impact is small as the vast majority of these projects are still in the research phase. The socio-economic impact is also slightly blurred as there

are still hurdles and ethical issues to be overcome with research into medicinal artificial intelligence such as the morality of allowing companies access to sensitive health data, the problems that stem from using healthcare data that is not representative of the whole population to train a model and the question of legal responsibility if a computer makes a mistake. However, despite all this the overall current socio-economic impact of artificial intelligence systems in medicine is positive if limited by the small number of systems that are actually used.

Abbreviations

Here is a list of commonly used abbreviations.

Abbreviation	Meaning
AI / ai	Artificial Intelligence.
FDA	Food and Drug Administration (referring to the Food and Drug Administration of the United States of America).
ML	Machine Learning (this refers to a subset of artificial intelligence where computers learn and adapt over time. A typical example of a structure used by machine learning techniques is a neural network (see paper for further details on neural networks)).
MRI	Magnetic resonance imaging.
CT	Computerised tomography.
USA / US	The United States of America
NHS	National Health Service (referring to the National Health Service of the United Kingdom).
NLP	Natural language processing.
AF	Atrial Fibrillation
DP	Decimal Place
UK	The United Kingdom of Great Britain and Northern Ireland
EU	European Union

Introduction

Artificial intelligence is a cutting-edge technology at the forefront of research and one that is predicted to have massive socio-economic ramifications in the near future. It is estimated that AI will lead to between 359.6 and 773.2 billion dollars of economic growth in the next five years [1], with a believed economic impact of $15.2 trillion by 2030 [2]. However, commercial adoption of artificial intelligence is yet to be widespread with there being a rather binary divide of industries that do and do not use it, with industries such as healthcare being prominent early adopters of AI [3]. Artificial intelligence raises a whole host of ethical issues which have serious social implications [4]. This has led many to question whether the economic gains artificial intelligence facilitates are cancelled out by the societal drawbacks that come with it, such as the furthering of inequity and reduced accountability due to the uncertain legal status of artificial intelligence systems. Equally so, one may argue that the artificial intelligence may instigate positive social change, with the technology boosting efficiency of workers, leading to general economic prosperity, and providing real benefits such as more reliable health care.

In this paper, the current socio-economic impact of artificial intelligence systems will be assessed in a field at the heart of AI innovation: healthcare. In doing so, one can analyse the current effectiveness of AI in the field which may be used by others to determine the suitability of more widespread usage of the technology in medicine.

Definitions

Artificial intelligence is defined as "the study and development of computer systems that can copy intelligent human behaviour" [5]. As such, when this dissertation refers to artificial intelligence systems, it is referring to computer systems that are able to perform tasks comparably or better than humans. This encompasses machine learning and deep learning systems but also covers more traditional computer programs such as rule and knowledge-based systems [106] if these systems are able to perform a task to the aforementioned standard.

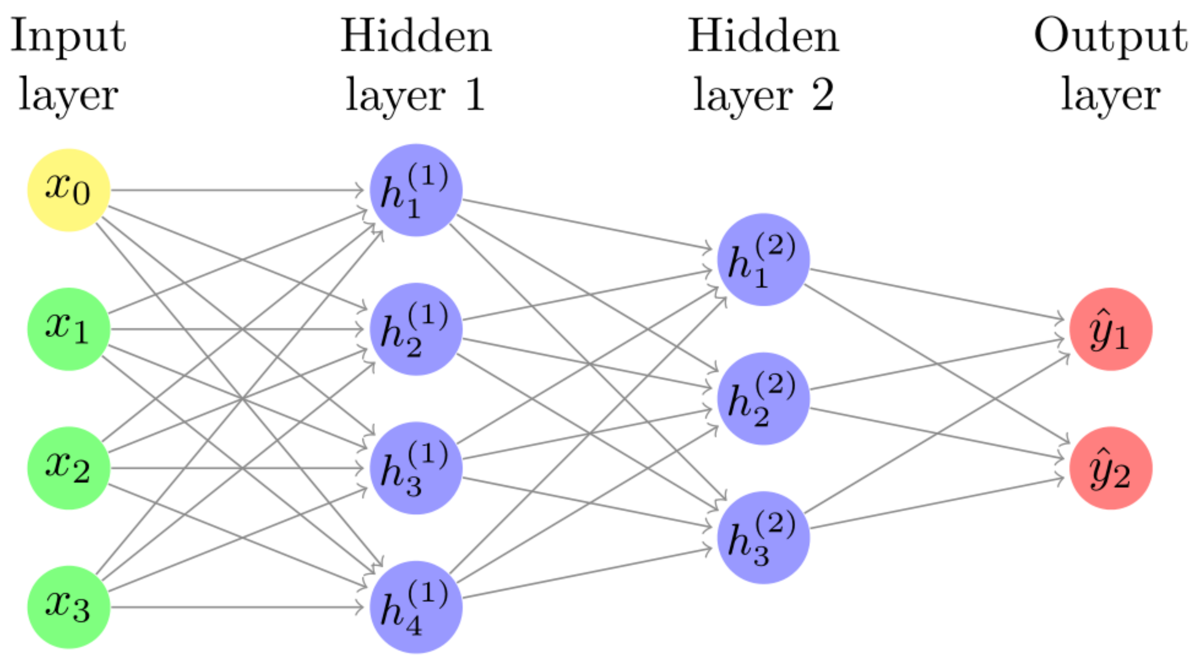
Medicine refers to the healthcare sector. Specifically, this essay defines medicine as:" the science or practice of the diagnosis, treatment, and prevention of disease" [80]. Hence, in this paper the socio-economic impact of AI in the healthcare / medicine sector will be evaluated.

This paper attempts to evaluate socio-economic impact, meaning that this paper will consider both the financial and social implications of artificial intelligence.

This paper is attempting to evaluate the overall socio-economic impact that artificial intelligence systems in medicine have on society and whether this is a net positive or negative one. This leads us to the question that the dissertation is attempting to answer: "To what extent do the use of artificial intelligence systems in medicine contribute to the socio-economic condition of society?".

Overview of Artificial Intelligence

Artificial intelligence can take many different forms, and can be as simple as human beings simply encoding knowledge into a computer [11] (e.g. if symptom x and y are present, probability of z disease is ø percent) or as complex as a self-learning system that mimics the structure of the brain and evolutionary mechanisms [81]. One such highly potent artificial intelligence technique is known as deep learning which makes use of a structure known as a neural network [86]. A neural network can be thought of as a highly complex non-linear, multi-variable function that consists of a large number of neurons that are used to convert input data into some form of output. Through mathematical techniques such as gradient descent [9], neural network can learn knowledge in a similar fashion to the brain and apply this knowledge to examples it has never seen before. To achieve this, labelled data sets are often required [107] (although there are also methods to train an AI without labelled data [10] [82] [90] [89]) with the accuracy and size of the dataset impacting the AI's overall accuracy [94] [95]. The ability of computer systems to learn in a humanlike fashion leads to a versatile range of applications, with deep learning systems being extremely competent with tasks such as image recognition [8] and natural language processing [7] [92].



[Figure 1 - Neural Network architecture. Image from: https://tex.stackexchange.com/questions/153957/drawing-neural-network-with-tikz (accessed February 19th 2020)]

[Figure 4 – Visualisation of a trained neural network recognising handwritten digits. This one has a success rate of 98.99%, although similar systems can achieve success rates of over 99.7% [20]. This tool and network were coded by Philip Mortimer (the author of this dissertation). https://github.com/philipmortimer/MNIST-AI (accessed 13th July 2020).]

Prevalence of Artificial Intelligence Systems in Medicine

Artificial intelligence in healthcare is a very new development and as such, there is a lot of research being conducted in the field but far fewer approved and implemented uses as systems have to be proven to function to a high standard. In a 2016 survey ([3]), health was found to be a leading industry in terms of AI implementation, although its usage was still assessed to between exploratory and experimental. AI seems to hold incredible promise in the field, particularly in natural language processing [84] and image recognition [83] [103]. Jiang et al. ([14]) summarise a few instances where machine learning techniques have been applied in healthcare to great success. The use of AI to identify and aid treatment of strokes is highlighted and shows the high accuracy rates are achieved, with one model by Similarly, Maninini et al. ([15]) able to accurately categorise different pathological gaits (stroke patients typically have identifiable gait patterns) 90.5% of the time. Jiang et al. however highlight that there are still large hurdles to the successful implementation of these techniques. For example, the American Food and Drug Administration (FDA) approved the first clinical based machine learning software in 2017 [16], showing just how recent advances in the field are. This also highlights the relatively low amount of current use of AI in healthcare. Despite this, there has certainly been a notable increase in the number of AI applications that have made it past the research phase and have been approved by the FDA. Figures 2 and 3 show that in January of 2019 there were 14 approved uses, but this had risen to 26 by April of that year. This rapid growth in actual implementation, shows that the field is very much translating lots of research into approved products. Whilst this number is small, there is clearly a tangible socio-economic impact that these applications have, particularly as many of these applications have to be proven to perform at least as well as humans to obtain approval. There is also a wealth of research occurring in the field [93], which can also be evaluated as there are also socio-economic implications stemming from such research projects, even if they are not as important as those from systems currently in use.

Table 2 | FDA AI approvals are accelerating

Company	FDA Approval	Indication
Apple	September 2018	Atrial fibrillation detection
Aidoc	August 2018	CT brain bleed diagnosis
iCAD	August 2018	Breast density via mammography
Zebra Medical	July 2018	Coronary calcium scoring
Bay Labs	June 2018	Echocardiogram EF determination
Neural Analytics	May 2018	Device for paramedic stroke diagnosis
IDx	April 2018	Diabetic retinopathy diagnosis
Icometrix	April 2018	MRI brain interpretation
Imagen	March 2018	X-ray wrist fracture diagnosis
Viz.ai	February 2018	CT stroke diagnosis
Arterys	February 2018	Liver and lung cancer (MRI, CT) diagnosis
MaxQ-AI	January 2018	CT brain bleed diagnosis
Alivecor	November 2017	Atrial fibrillation detection via Apple Watch
Arterys	January 2017	MRI heart interpretation

[Figure 2. FDA AI approvals as of January 2019. https://twitter.com/EricTopol/status/1119683505603006469 (accessed on 8th June 2020)] Note: others have highlighted other AI applications that may also qualify for this list. Some of these examples can be found in the comments of the original post. It is also important to note that some other AI systems are used in healthcare that do not require FDA approval but most of these fall outside of the scope of the dissertation as they do not directly satisfy this essay's definition of medicine. The number of these systems that have already been implemented is also low [98].

Quantib	Medical imaging (MRI)
CureMetrix	Medical imaging (Mammography)
BrainScope	Concussion Assessment (multi-modal)
Apple	Atrial fibrillation detection
Aidoc	Medical imaging (CT scan)
iCAD	Medical imaging (Mammography)
Zebra Medical	Medical imaging (CT scan)
Bay Labs	Medical imaging (echocardiogram)
Neural Analytics	Device for Paramedic stroke Diagnosis
IDx	Diabetic retinopathy diagnosis
Icometrix	Medical imaging (MRI)
Imagen	Medical imaging (X-ray)
Viz.ai	Medical imaging (CT scan)
Arterys	Medical imaging (MRI and CT)
MaxQ-AI	Medical imaging (CT)
Alivecor	Atrial fibrillation detection
DreaMed	Diabetes treatment decision
Arterys	Medical imaging (MRI)
Empatica	Warning of seizure risk
Subtle Medical	Medical imaging
Cognoa	Autism diagnosis
Healthy.io	Urinary tract infection diagnosis
Excel Medical	Remote monitoring
FibriCheck	Atrial fibrillation detection
ScreenPoint Medical	Medical imaging (Mammography)
SyncThink	Eye movement disorders

[Figure 3. FDA AI approvals as of April 20th, 2019. https://twitter.com/EricTopol/status/1119683505603006469 (accessed on 8th June 2020)] Note: others have highlighted other AI applications that may also qualify for this list. Some of these examples can be found in the comments of the original post. It is also important to note that some other AI systems are used in healthcare that do not require FDA approval but most of these fall outside of the scope of the dissertation as they do not directly satisfy this essay's definition of medicine. The number of these systems that have already been implemented is also low [98].

Socio-economic Impact of Artificial Intelligence Systems in Medicine

As previously mentioned, AI systems in healthcare can be broadly broken down into two categories: currently implemented and currently being researched. The vast majority of systems are in the latter category as large hurdles still remain for wide-spread implementation [19] [91] [98].

Many industry stakeholders are still hesitant to invest as they feel that the financial benefits of the technology are yet to be proven [19], which is argued to lead to a cycle of a lack of investment which in turn means that there is little evidence to show the financial benefits of the technique. Fortunately, it appears that the industry is becoming more aware of the capabilities of AI as numerous studies prove the potential for AI to cut costs and make treatment more effective. This idea that research leads to investment shows that researching AI applications in medicine does have a socio-economic impact as some of the systems being researched will make it to market, and others will encourage further investment in the field.

Research into Artificial Intelligence Systems that could be used in Healthcare

Deep learning systems are highly adept at analysing large amounts of data, particularly data that is labelled. As such, one of the most common applications of AI is image recognition [103] [104], with advanced techniques (such as convolutional neural networks) enabling computers to recognise images correctly as often as 99.77% of the time [20]. The healthcare sector makes use of a wide range of imaging techniques such as x-ray scans, MRI (magnetic resonance imaging) scans and CT (computerised tomography) scans which are invaluable in both diagnosis and prevention of disease. NHS England performed a total of 40.2 million imaging tests over a twelve-month period spanning from July 2015 to June 2016. However, response time was not flawless, with patients who received an MRI scan waiting an average of three days after the procedure for the report. Furthermore, time between request for a scan and scan being carried out was highly variable ranging from the same day to a wait of 21 days [21]. Software that is able to automate or speed up the process of analysis of scans would both reduce waiting times as well as time for reports to be released [102]. As such, artificial intelligence holds great promise in freeing up doctors to perform other tasks as well as improve diagnosis accuracy.

One such recent breakthrough is an Artificial intelligence system developed to analyse mammograms in order to determine whether a person had breast cancer. The computer was trained on scans from nearly 29,000 women [23] and was found to outperform a radiologist in the UK, with a 1.2 percent reduction in false positives and a 2.7 percent reduction in false negatives. It performed even better in comparison to radiologists in the USA with a reduction of 5.7 percent in false positives and 9.4 percent in false negatives [22]. Furthermore, the system was able to perform just as effectively as two radiologists working together. Currently, a mammogram is analysed by two radiologists and if disagreement about the results exists, a third radiologist is consulted [23]. Reading scans is “time-consuming work” ([23]) and one can typically expect to wait up to two weeks to receive the results from a screening [25]. This is highly promising as the AI could be used as an alternative to a second radiologist which would speed up the processing of results (which is highly important as early identification of breast cancer dramatically increases survival chances [26], even though two weeks is a relatively small period of time) and would help address the shortage of radiologists which is estimated to be over 1000 [23]. In this instance, there is large economic potential as it speeds up the vital work of radiologists and ultimately makes them more cost-effective to the taxpayer. Increased diagnosis accuracy has massive economic value also as increased chances of survival allow patients to contribute economically to society (not to mention the massive social benefit that comes with the saving of a life). However, this paper is considering the current socio-economic impact which is much more limited. Short term economic impact is arguably negative as research of this manner does come at a price. However, research of this sort is an investment and given the extremely promising results, the long term-economic impact of this will be highly positive as it will improve accuracy of results which will lead to fewer deaths. Computers have the advantage of performing consistently and not getting tired and thus they can provide faster and more consistent analysis [96]. Therefore, even if no changes were made to the system apart from using the computer in addition to the radiologists, accuracy and confidence in results could be increased. The AI would also enable doctors to prioritise more urgent mammograms first. Therefore, the overall economic impact is positive but limited as it will encourage more research and eventually lead to better treatment of diseased albeit at considerable current cost. There are also social considerations, not the least the implications of granting access to 29,000 scans. Naturally, some feel uncomfortable at the prospect of handing over intimate data to Google and some argue that patients have a right for data of that sort to remain private. However, I would argue that appropriate safeguarding measures have been taking as the data has been anonymised [23] which sufficiently addresses these fears. However, the precedent of granting access to sensitive data for research to third party organisations is nonetheless one that many will be uncomfortable with, particularly as they may have no say in the matter. Despite this, the societal benefits of improved breast cancer screening processes drastically outweigh such risks which can be managed by steps such as anonymisation. Studies like these provide a great social benefit as they encourage more investment in research into technology such as this, which will lead to improved healthcare treatment. Therefore, I conclude that research into AI does in fact have a net positive impact as it facilitates improved healthcare conditions, although there are certainly important negative aspects that need to be considered.

This study into breast cancer diagnosis is typical of a large range of other studies which all work on use of an AI to replicate a doctor analysing a scan [105]. Other imaging AI research that has comparable (although not always as successful) results includes detecting liver disease [27], treatment of muscular sclerosis [28], and aiding treatment of cardiovascular disease [29]. As much of the research into AI in medicine is of similar nature (namely enhancement of imaging), one can conclude that the socio-economic implications of research into radiology are broadly applicable to wider research. With over $500 million invested in Medical Imaging AI companies across a six-month period in 2018 [30], there is certainly a large level of investment in the field. This naturally has negative short-term economic implications as such large amounts of money are being invested into a technology which is unlikely to return this investment in a short period of time. The long-term economic impact is far less clear as the technology has massive promise, however there are still significant hurdles to overcome for wide-spread implementation to become realistic. Investment in AI research undoubtedly makes the implementation of such technology in medicine a much more realistic prospect in the future and will work to overcome the challenges it faces. If AI can overcome certain hurdles, the economic impact will ultimately be positive as it will enhance medical practice. A review of 69 studies published in the Lancelet found that deep learning models were able to perform at a level similar to healthcare professionals [37], showing that if these hurdles can be overcome, the potential socio-economic rewards are large-scale and obtainable in the foreseeable future.

Problems with Implementation of Artificial Intelligence Systems in Healthcare

Artificial Intelligence systems face enormous hurdles before they can become widespread in the healthcare field [85]. All AI needs to be approved before it is used in medicine by agencies such as the FDA, which require high standards to be met before anything can be used in such a field.

There are difficult legal issues to be evaluated when using computer systems [100]. Medical error is a serious issue, with a John Hopkins University study assessing medical error to be the third highest cause of death in the US in 2016, with over 250,000 US deaths a year attributed to it [31]. This leads to a large number of medical malpractice cases per year [32]. However, who is legally responsible if a computer system makes a poor decision? This is a complex and still somewhat unclear legal question, although AI systems that are used to enhance the strength of a decision of a doctor as opposed to replacing it are believed to be acceptable. The FDA appears to be accepting of machine learning software used for a “non-serious situation” ([33]) and is open to use of AI in all areas of the field, but the simple truth is that most AI systems in current use are used to support Doctors as opposed to replacing them. However, AI is capable of replacing doctors in certain fields. If the AI then makes a mistake that results in the death of a patient, who should be held accountable? One of the flaws of AI, much like humans, is that it is almost certain to act imperfectly. It may even perform better than humans, but it is much harder to hold a computer accountable for malpractice. At the current time point, legislation and guidance is unclear as to who may be accountable and varies from country to country. However, AI is not yet at a level, where this is too much of a problem as it is typically only used to help medical professionals, not to replace them.

Huss, M ([34]) points out that ML AI is limited in its ability to aid doctors as self-learning models that make use of architecture such as neural networks can often be thought of as a black box. This is because the model is attempting to extract meaningful information from a given data set so that it may produce an output (such as the probability that a given patient has breast cancer). However, such models often have millions of variables and it is often impossible to determine what these models are looking for. Thus, AI lacks the expressiveness and ability to justify a decision which is of great importance to a doctor. Self-learning models are limited by the amount of and quality of data available to them and as such may be deeply flawed if the data set, they are trained on does not include a wide-enough range of information. This is a substantial flaw and has serious social implications as it may mean that members of the public who have a rarer instance of a disease (that the AI may not have learned about) will be adversely impacted. Furthermore, the black box nature of such systems makes it harder for doctors to trust them, particularly when a doctor disagrees with the computer [100]. It is important to note that the article by Huss has a potential conflict of interest, as Huss is developing a platform to address some of the issues he describes in the article (although this would of course make him knowledgeable on the subject). Despite this, the issues described is a common problem with AI and the debate around the usability of black-box model is certainly there. A piece in the British Medical Journal ([36]) concludes that AI has the ability to radically improve diagnosis but that the lack of clarity around the decision making process is a large weakness which limits the trust that doctors will assign to it. Anna Goldberg, a scientist in genetics, argues that the priority is to build the “best system” and then develop tools that interpret what the AI is doing[35], although no such tool has been properly realised yet. However, Katherine Andriole points out that machine learning models are predominantly black boxes but that use of technology such as a saliency map (which shows which features the model is analysing [102]) and analysis of the network can make it more understandable [98]. However, Andriole goes to on to stress that the black box nature is still a primarily unsolved problem, particularly as the FDA says this issue still needs to be overcome [98]. Given all this, I feel that if one can develop an AI that is of sufficient accuracy and convey this accuracy to the physicians using it, they should be confident in using the tool even if they may not fully understand how it operates. However, the black box nature of AI certainly has an overall negative social impact as it leads to distrust and also makes it much harder to detect AI systems that may be failing to properly consider important factors. But this on it is own is not a real hurdle to implementation, particularly when considering the accuracy of various healthcare research projects [37] and the fact that some black box systems have already been implemented into healthcare. Furthermore, AI is far from the only black box in healthcare with many tests displaying similar if not lower accuracy rates such as the RT-PCR test for COVID-19 (which has a believed false negative rate of between 2 and 29% [40] [41], although it must be noted that the study has not been peer reviewed and that this figure is still far from certain).

Whilst there are certainly are hurdles to be overcome as medical AI transitions from primarily research to actual implementation, they can certainly be overcome with adequate effort. This is proven by the fact that numerous AI applications have already been approved by the FDA for use, including the first medical imaging AI that detects lesions [38]. This along with statements by the FDA [39] show a willingness to work with researchers to make the transition from research to approval although the lack of explicit guidance is still undoubtedly a hurdle. Therefore, I conclude that whilst challenges such as legal accountability of AI are still prevalent issues, they do not significantly prevent a strong deep learning model transitioning from development to use (as evidenced by the approval of the lesion detection AI [38]). This makes the socio-economic impact of AI research in healthcare far more positive as it shows that research into the field can directly translate into use even if the process may be arduous, slightly ambiguous, and sometimes lacking in precedent. Thus, research can in many cases simply be viewed as a foundation that establishes the viability of a system and given the high viability of many systems [37], it can be concluded that research leads to implemented AI systems, which have a positive socio-economic impact.

Artificial Intelligence Systems currently used in Healthcare

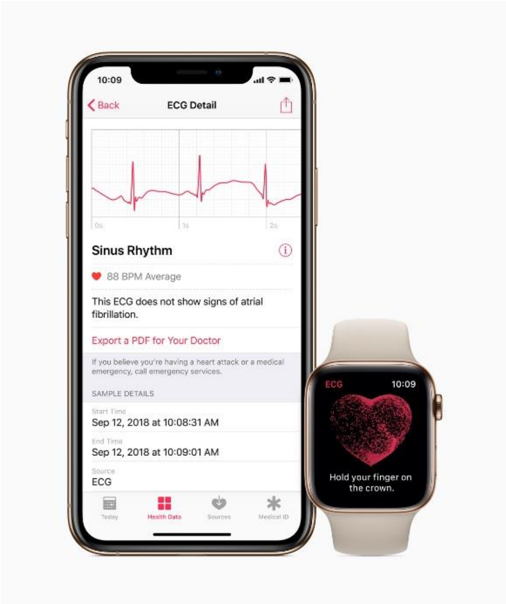
Although, research projects do have a socio-economic impact, AI systems that are used in healthcare contribute much more to the current wellbeing of society as they directly impact patients. As previously discussed, there are few AI products currently approved for use by the FDA (see figure 2 and figure 3), however these systems may be limited in number but still certainly have a substantial impact on healthcare. In addition, artificial intelligence systems from other fields are also making a tangible impact in the healthcare sector. Typically, AI in healthcare is thought of as bespoke software used to target an aspect of healthcare like diagnosis, however more conventional AI such as natural language processing software can also be applied in the field of care [42] and medicine [43], which benefits many groups of people, such as those with a disability. These systems include voice assistants such as Amazon’s Alexa and undoubtedly have large positive social impact as they often enhance the mental health of the lonely in places like care homes. However, systems like this fall outside of the scope of this essay as they are not medicinally useful (as they are not used to treat, prevent or diagnose a disease, the definition of medicine used for this essay) although they may of course still have a positive impact to many in the same way a better Facebook algorithm may enhance the wellbeing of an individual.

Artificial Intelligence Systems currently used in Medicine that Perform Specialist Tasks

Atrial fibrillation occurs when “electrical impulses fire off from different places in the top chambers of the heart (the atria) in a disorganised way” [45]. This results in arrhythmia (an irregular heartbeat) and can lead suffers to feeling tired, short of breath or faint, although many are asymptomatic [45] [46]. However, atrial fibrillation is difficult to detect leaving an estimated 13.1% of atrial fibrillation cases undetected [47], which is significant given the prevalence of the condition which is thought to be present in 5 million US citizens [47] and 1.2 million British citizens [48]. This means that there are an estimated 700,000 people in the USA who are unaware that they have the condition [47], which is massively important given that people with untreated AF are five times more at risk of a stoke [48] than the average population and that 31.2% of strokes are associated with AF (although this number is variable as it is based on a sample size 568 stroke patients and as the cause of stroke may be another factor such as a poor diet) [55]. The issue also has massive economic ramifications as undiagnosed AF in the USA is estimated to have a cost of $3.1 billion [49], with a study into the issue concluding “strategies to identify and treat patients with undiagnosed AF could lead to sizable reductions in stroke sequelae and associated costs” [49]. Given this, there is massive economic and medicinal benefit to detect AF in cases where people may be unaware that they suffer from the condition, as through diagnosis and subsequent treatment, the risk of a stroke can be reduced by 64% [50]. However, it is difficult to devise a strategy to identify patients with undiagnosed AF even through measures such as randomised check-ups (which in themselves would be cost ineffective) would fail to catch many cases as many sufferers have “AF episodes that come and go” (known as paroxysmal AF) [51].

It is estimated that over 70 million people use, or have used an Apple Smart Watch, since its initial release in 2015, with this number anticipated to continue to rise rapidly [53][52]. Included in the device is a heart rate monitor which has an accuracy of 99.9% when compared to a professional clinical pulse oximeter (the Onyx Vantage 9590), which means that the Apple Watch is effectively as accurate as a clinical device [54]. Given this much more frequent and highly accurate access to a person’s heartbeat, Apple looked into the possibility of using this to detect AF. Through the training of a deep (many layered) neural network from data provided by 9750 participants enrolled in the Heart eHealth Study and 51 patients undergoing cardioversion (which resulted in over 139 million data items), an AI was developed that had a sensitivity (true positive rate) of 98.04% and a specificity (true negative rate) of 90.02% [56] [57]. Apple obtained FDA approval and the Apple Watch now has the ability to detect AF in certain countries [58] [59] [Figure 2] [Figure 3]. This technology is not unique to Apple as both FibriCheck [61] [Figure 3] and Alivecor [60] [Figure 2] [Figure 3] also have FDA approved AF detection technologies. FibriCheck’s software is particularly unique in the sense that it detects AF using just a smartphone and no external accessories (although this does of course come with the down side that patient has to actively choose to record their heart rate, unlike a watch which may be worn more regularly). Given the large market of smartphones, with roughly 1.5 billion sold in 2019 alone [62], and smart watches [53] [53] such as the Apple Watch, there is undoubtedly the potential to detect a large number of the asymptomatic AF sufferers and thus save lives.

In order to determine the impact of AI systems that detect AF, this paper will analyse The Apple Heart Study. The Apple Heart Study was carried out across 419,093 participants who were over the age of 22 and had not been diagnosed with AF and was designed to determine the proportion of patients who received irregular pulse notifications from their Apple Watch that were subsequently diagnosed with AF [63]. The large sample size means there can be a high confidence in the results obtained, although the study design is limited by a few factors [88]. The study failed to meet its target of 500,000 participants which in itself is of little consequence as the sample size is still extremely large. It also failed to meet its target of 75,000 participants over the age of 65 which is of greater consequence as AF is much more prevalent in the elderly. However, one could argue that this may be somewhat of a reflection of the younger demographic of the Apple Watch and highlights more of a flaw with the use of the technology as a means of diagnosing AF as opposed to with the study itself. The accuracy of the study is limited by its reliance on the self-reporting of patients [64] which would reduce confidence somewhat, however this too is likely to be negated by the large sample size. As one had to register for the study digitally, one could argue that the demographic may be skewed slightly towards those who are digitally literature.



[Figure 6. An example of the health app tracking the heart. https://www.apple.com/newsroom/2018/12/ecg-app-and-irregular-heart-rhythm-notification-available-today-on-apple-watch/ (accessed on 14th July 2020).]	[Figure 7. An example of an irregular pulse notification. https://www.apple.com/newsroom/2018/12/ecg-app-and-irregular-heart-rhythm-notification-available-today-on-apple-watch/ (accessed on 14th July 2020).]

Overall, 0.52% of participants received irregular pulse notifications, with the pulse detection algorithm being found to have a 71 percent positive predictive value. However, of the people who received an irregular pulse notification and followed up, only 34% were found to have atrial fibrillation. The study points out that this does not mean that only 34% had AF, but that many of them may have had paroxysmal AF that simply was not present during the follow up. Only 0.16% percent of participants under the age of 40 received a notification, whilst just over 3% of participants over the age of 65 received one, highlighting just how much more prevalent the condition is in the elderly. Given the estimate of 700,000 [47] cases of undiagnosed AF in the USA in 2009, one can conclude that roughly 0.2285% [65] have undiagnosed AF. In the study, 0.1768% (34% of 0.52%) of participants were found to have AF, which is roughly 77% of the number of expected AF cases. This should serve as an upper estimate for the capability of the smart watch as its accuracy is negated by the fact that many people are unlikely to trust the Watch's diagnosis which may lead them to ignore notifications. This issue can be circumvented by education, but it is important to note given that only 57% of notification receivers sought healthcare. [66] [64] [67] [Figure 5]

Given the estimated economic burden of $3.1 billion [49], if the Apple Watch has the potential to identify 77% of these cases, it has the potential to drastically reduce the economic burden of AF by around 2.4 billion dollars in the US alone. If one applies this to the rest of the world (which will lead to a very rough figure estimate), one can conclude that the Apple Watch has the potential to save 55.9 billion dollars (which is extremely unrealistic as it assumes that every single person in the world would wear a watch). Of course, the actually impact of the watch is significantly lower due to the younger demographic, lack of trust in the in the technology as well as the fact that the technology comes with a £300 price tag. However, if we take the estimate of 70 million [53] [52] apple watch owners and combine it with the result of the study, the Apple Watch would save 548 million dollars (although not all 70 million are active users of the watch). And this number would only grow with increased education, sales of the Apple Watch (as each sale, on average, saves 7.8 dollars due to AF diagnosis which assumes that all watch wearers have the AF diagnosis feature turned on) and broadening of the demographic (i.e. more sales to older people). [Figure 5]

Percent of population with undiagnosed AF in 2009 in USA = number with people with undiagnosed AF in USA / population of USA in 2009 = 700,000 [47] / 306,307,565 [65] = 0.228528373% (rounded to 9 DP)

Percent of Apple Heart Study with undiagnosed AF that was detected = percent of irregular pulse notifications * percent of AF cases diagnosed following irregular pulse notifications = 0.52% [67] * 0.34% [67] = 0.1768%

Percent of cases that Heart Study identified = Percent of Apple Heart Study with undiagnosed AF that was detected / Percent of population with undiagnosed AF in 2009 in USA = 0.1768 / ((700,000 [47]) / 306,307,565 [65]) = 77.3645393% (7 DP)

Potential (i.e. maximum) Economic Burden Reduction of Apple Watch through AF diagnosis in US (assuming all US citizens have an apple watch, which they do not) = economic impact of undiagnosed AF in US * Percent of cases that Heart Study identified = 3.1*10^9 [49]* (0.1768 / ((700,000 [47] * 100) / 306,307,565 [65])) = $2,398,300,718 (Note: this is a lower estimate as they 3.1 billion has not been adjusted for inflation from 2015)

Potential (i.e. maximum) Economic Burden Reduction of Apple Watch through AF diagnosis in the World (assuming all citizens have an apple watch, which they do not) = Potential (i.e. maximum) Economic Burden Reduction of Apple Watch through AF diagnosis in US (assuming all US citizens have an apple watch, which they do not) * (population of world / population of USA) = 2,398,300,718 * (7,763,035,303 [68] / 332,865,306 [69]) = $55,932,813,677 (rounded to nearest whole number)

Current economic savings per Apple Watch due to AF diagnosis = Percent of Apple Heart Study with undiagnosed AF that was detected * (economic impact of undiagnosed AF in US / number of people with undiagnosed AF in US) = 0.001768 * ((3.1*10^9 [49])/700,000 [47])= $7.82971429 (rounded to 8 DP)

Current economic savings caused by Apple Watch's ability to diagnose AF (assuming all people who have ever purchased an Apple Watch still wear it regularly) = Current economic savings per Apple Watch due to AF diagnosis * estimated number of apple watch sales = (0.001768 * ((3.1*10^9 [49])/700,000 [47])) * 70,000,000 [53] [52] = $548,080,000

[Figure 5. Mathematics performed by author of the dissertation (Philip Mortimer). Last updated: 6th July 2020.]

Clearly, the current economic impact of the Apple Watch is large although not as large as $548 million (although still in the many millions) as the current user base of the Apple Watch is smaller than this, as the number of users who seek healthcare following a notification is only 57% and as the Apple Watch AF detection software is region specific (i.e. it has not been approved by all of the relevant global health agencies, although it has been approved by the FDA [87]). It is also limited by the fact that this feature is automatically turned off and needs to be turned on by the user. However, all of these issues can be counteracted through education and as such are not serious hurdles in the long term. Of course, the economic impact extends far beyond just the Apple Watch as this technology has the potential to be used by all wearables as well as mobile phones (although mobile phone AF detection is less advanced as the device is not taking continuous measurements) which covers about half of the world [70]. The social impact is naturally large too as reduction in healthcare expenditure is advantageous as is the saving of lives. However, the Apple Watch will help save lives as treatment of AF is believed to reduce stroke risk by between a half and two-thirds [71] especially given that 28% of stroke victims die within a month of the stroke, increasing to 41% after a year [72]. As a result of all this, AI is saving many hundreds of lives and millions of dollars through AF diagnosis. And as adoption of technology such as smart watches and clinical approvals increase, the number of lives and economic value provided by the software will increase. And as many point out, pulse detection and analysis algorithms incorporation with technology can be extended to a range of heart diseases. However, the benefit is greater than that still as this technology is already used in fitness tracking software which improves general wellbeing as it inspires and enables greater fitness levels.

A large amount of other approved artificial intelligence in medicine revolves around medical imaging with 13 out of the 26 FDA approvals in Figure 3 being used for imaging. At the moment, most of these imaging tools analyse images in a similar way to the aforementioned mammography AI with the key difference being that the systems are primarily used in order to strengthen the decision of a doctor as well as prioritise human analysis of those that are more likely to be at immediate risk. As such they undoubtedly do increase the accuracy of medical advice and enable faster treatment for those who require it. There are far fewer approved diagnostic imaging devices, however one such approval is for the diagnosis of diabetic retinopathy which the FDA has approved to be utilised without a specialist verifying results [73] with the system managing a roughly 90% accuracy [74]. This naturally has a large impact as over 24,000 cases of blindness each year in the USA are attributed to the disease and as blindness is believed to cost $5.5 billion in annual medical expenses in the US as well as 209000 quality-adjusted life years in the USA [75]. Through accurate screening using AI, doctors will be empowered to reduce the number of blindness cases caused by diabetic retinopathy which will help massively economically as blind people often struggle to contribute to the economy and socially as blindness is a life changing impairment. However, if we look at AI in medicine as a whole, we see that diagnosis AI is rare although massively useful in the fields it is used in.

Conclusion

Artificial intelligence in medicine has massive potential with the qualities of deep learning models suited for medicinal tasks such as image analysis to determine whether a disease is present or not. This is reflected in the large amount of research in the field with a vast range of software that performs better than or as well as doctors in a wide range of medicinal fields. However, whilst the promise of AI is clear, most research is yet to be translated into implementation for numerous reasons. As the field is so new there is no precedent as to whether the creator of the software could potentially be held accountable for incorrect decisions. The FDA and other bodies appear to be willing to provide guidance and approve AI in the field, however companies are hesitant to invest in such a new field which still needs to answer some questions. In order to assess the overall current socio-economic contribution of medicinal artificial intelligence, one must analyse both currently implemented AI and research in the area.

Research into AI in healthcare is wide-ranging and has undoubtedly proven its ability in many fields to perform as competently as a doctor without the drawback's doctors have such as getting tired. Undoubtedly much of this research will be implemented in the next few decades, leading to massive socio-economic gains. One could therefore argue that research facilitates implementation of AI and does possess worth as it is required in order to implement these lifesaving algorithms. For example, the model that is able to diagnose breast cancer more accurately than a physician has the ability to save countless lives and through research, such technology is able to gain approval for medicinal use as it is demonstrating its ability. One can imagine the benefits of research to be similar to the benefits of investing in development of a vaccine, in the immediate short term it may be costly, but in the long term it has the potential to save countless lives. Therefore, there is a slightly negative short-term economic impact of research as it is costly, but it has a massively positive long-term impact as it leads to implementation of advanced systems as well as the improvement of such algorithms through collaborative research. Research also proves the potential of AI, which encourages others to invest in the technology which will ultimately lead to a better healthcare system. All of the currently implemented AI systems were made possible through research which shows that it does have a current socio-economic impact, although this is limited due to the small number of approved use cases. There are also detrimental social considerations as advanced AI models depend on large amounts of data to be trained. This leads to some privacy concerns as sensitive data such as breast cancer scans are often divulged to third party companies such as Google without the patients knowledge, although mitigation strategies such as anonymisation of the data have quelled these fears somewhat. There is also a fear that patients who display atypical symptoms of a certain disease may be misdiagnosed by systems as the black box nature of AI makes it near impossible to discern what an AI actually uses to predict whether an individual has a disease. For example, Amazon once built an AI to filter through resumes to determine strong applicants and was trained using the employment records of the company. As these records were of predominantly male employees, the AI perpetuated this bias and interpreted this to mean that a female applicant was by default a far weaker applicant than a male one [76]. There are fears that this premise may translate over in the technology if designers are not careful as it could interpret, for example, that somebody may be more likely to have a certain illness because of their race. For example, if a disproportionately low number of patients from a certain ethnicity are represented in the training data this bias may carry across to implementation [100] [103]. This may further the racial inequity present in healthcare and quality of healthcare [77] [78] [79]. However, there is little evidence to show that these fears are warranted yet, partially because, so few algorithms have been implemented and hence tested on such a large scale. This is also an issue that can be monitored against by limiting irrelevant data and ensuring sample data is large and representative. Given all these considerations, this paper concludes that research into medicinal AI is massively advantageous in the long term. However, its current socio-economic impact is more limited. This paper argues that the current economic impact is limited as research is very expensive, however it is argued that the impact is still positive as it has led to a very small number of implementations, but that these implementations have a large positive economic impact right now. The societal impact is more mixed as research has led to systems that do save many lives and enhance the quality of life for many people. However, ethical concerns doubtlessly remain about the use of computers to diagnose diseases with questions about who is responsible if such a system fails. There are also concerns both about the morality of companies accessing sensitive healthcare data as well as the quality of such data potentially leading to disaster such as the furthering of healthcare inequity due to things such as race. One could also counterpoint this by saying that this research will save possibly millions of lives, which is massively socially beneficial. It is also important to note that many of hurdles to widespread implementation can be overcome through legislation, governmental guidance on practices to maintain when developing AI systems in medicine and careful design of software. The long term social impact is highly positive, but the current socially impact is slightly negative as a result of these concerns (which will be overcome with time but persist at the moment), although it must be acknowledged that the research has led to a small range of systems that already make a large positive social difference.

Currently implemented AI systems in healthcare are clearly have a much larger overall current impact. For example, Apple Watch's atrial fibrillation detection feature saves an average of $7.8 per user who enables the feature in healthcare expenditure, which extrapolated over the roughly 70 million Apple Watch users leads to millions of dollars in healthcare treatment saved. It also leads to far fewer people suffering from strokes, therefore both saving lives and leading to lives of better quality. Similar AF algorithms are also used by other companies and allow for constant monitoring for the condition through smart watches and smart phones. One could also point out that whilst a minor drawback, it could potentially further the health divide between the rich and poor, although this isn't really a drawback as the overall quality of healthcare is improved, it is just, unfortunately, less accessible to those less well off. There a few approved AI systems that diagnose disease with human like accuracy that facilitate more effective screening for something like diabetic retinopathy as experts are not needed to analyse every scan initially. This means that a larger group of people who are at risk can be scanned and thus more cases are picked up. Around half of all AI currently approved by the FDA is used in imaging and is often used to strengthen a doctor's decision or to screen a wider pool of people more quickly, with the doctor then being able to target those who need professional analysis first. Each of the systems currently implemented has an incredible economic impact as they are highly effective in increasing quality of diagnosis and healthcare. And naturally, better diagnosis and treatment saves more lives and enhances the quality of life which often leads to reduced medical bills and facilitates more people to work and thus improve the economy. This is not to mention the fact that diagnosis of something such as AF or diabetic retinopathy leads to large reductions in medical fees. There are obvious and substantial social benefits in enhancing and saving thousands of lives which appear to come with little drawback apart from perhaps a lack of trust in these systems as well as the black box nature of these systems potentially leading to serious undetected mistakes. The current impact is primarily limited by the fact that there are so few systems currently implemented (with around 26 being approved for use by the FDA in April of 2019 [Figure 3]) although this number appears to be growing rapidly [Figure 2] [Figure 3] which suggests that AI adoption will be widespread in the healthcare sector in the coming decades. At the moment the socio-economic impact is large in the sectors that it is used in and contributes many millions of dollars (if not billions) and saves and improves the quality of thousands of lives in spite of the fact that is used so little. Its use is small relative to the size of the whole medicine sector, but in the small number of places where it is used, it is effective and makes enormous socio-economic contributions.

One can conclude that AI in healthcare has a sizeable and substantially positive socio-economic impact although it is severely limited by the fact that its implementation is not widespread and slightly limited by societal implications regarding the dependency on large amounts of sensitive data. The economic impact of AI in healthcare is overwhelmingly positive as it reduces healthcare expenditure and saves lives, which enables more people to contribute to the economic wellbeing of society. Although investment into research is expensive, this expenditure will be paid back many times over in the long term and has already been paid back with algorithms such as AF detection which save millions of dollars annually. Economic and social benefit are linked to each other as saving of lives and diagnosis of disease lead to positive societal and economic gains. As such, research and implementation are socially beneficial as they save and enhance the quality of lives. As a result of all this, this paper concludes that artificial intelligence systems in medicine have a large positive social and economic impact even if this is substantially limited by the fact that only a very small number of artificial intelligence systems are currently used in medicine. However, this just shows precisely how positive the socio-economic impact of each artificial intelligence system in medicine really is.

Bibliography

[1] Chen, N., Christensen, L., Gallagher, K., Mate, R. and Rafert, G., 2016. Global economic impacts associated with artificial intelligence. Analysis Group. Accessed on 4th June 2020. https://pdfs.semanticscholar.org/ebc7/3c75f7aba486751a20257a3134c777afd255.pdf

[2] Gillham, J., Rimmington, L., Dance, H., Verweij, G., Rao, A., Roberts, K. and Paich, M., 2018. The macroeconomic impact of artificial intelligence. PricewaterhouseCoopers. Accessed on 4th June 2020. https://www.pwc.co.uk/economic-services/assets/macroeconomic-impact-of-ai-technical-report-feb-18.pdf

[3] McCauley, D., 2016. Artificial intelligence in the real world: The business case takes shape. The Economist. Accessed on 4th June 2020. https://eiuperspectives.economist.com/sites/default/files/Artificial_intelligence_in_the_real_world_1.pdf

[4] Bossmann, J., 2016. Top 9 ethical issues in artificial intelligence. World Economic Forum. Accessed on 4th June 2020. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

[5] Oxford Learner's Dictionary. Accessed on 4th June 2020. https://www.oxfordlearnersdictionaries.com/definition/english/artificial-intelligence

[6] Cambridge Dictionary. Accessed on 4th June 2020. https://dictionary.cambridge.org/dictionary/english/data-analysis

[7] Jurafsky, D and Martin, J (2019). Neural Networks and Neural LanguageModels. https://web.stanford.edu/~jurafsky/slp3/7.pdf . Accessed 19th February 2020.

[8] Sanderson, G (2017). Neural Networks. https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi (accessed on 24th November 2019).

[9] Hallström, E., 2016. Backpropagation from the beginning. https://medium.com/@erikhallstrm/backpropagation-from-the-beginning-77356edf427d (accessed on 5th June 2020).

[10] Machado, G., 2016. ML Basics: supervised, unsupervised and reinforcement learning. https://medium.com/@machadogj/ml-basics-supervised-unsupervised-and-reinforcement-learning-b18108487c5a (accessed on 5th June 2020).

[11] Tricentis. AI Approaches Compared: Rule-Based Testing vs. Learning. https://www.tricentis.com/artificial-intelligence-software-testing/ai-approaches-rule-based-testing-vs-learning/ (accessed on 5th June 2020).

[12] Chui, M and Malhotra, S., 2018. AI adoption advances, but foundational barriers remain. https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain (accessed on 7th June 2020).

[13] Cam, A., Chui, M and Hall, B., 2019. Global AI Survey: AI proves its worth, but few scale impact. https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact (accessed on 7th June 2020).

[14] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H. and Wang, Y., 2017. Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology, 2(4), pp.230-243. https://svn.bmj.com/content/2/4/230.full (accessed on 8th June 2020).

[15] Mannini, A.; Trojaniello, D.; Cereatti, A.; Sabatini, A.M. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. Sensors 2016, 16, 134. https://www.mdpi.com/1424-8220/16/1/134/htm (accessed on 8th June 2020).

[16] Marr, B., 2017. First FDA Approval For Clinical Cloud-Based Deep Learning In Healthcare. https://www.forbes.com/sites/bernardmarr/2017/01/20/first-fda-approval-for-clinical-cloud-based-deep-learning-in-healthcare/#7ba305f9161c (accessed on 8th June 2020).

[17] Guttmann, A., 2018. Marketing-related use of artificial intelligence (AI) according to industry professionals worldwide as of January 2018, by region. https://www.statista.com/statistics/915372/marketing-related-use-artificial-intelligence-world/ (accessed on 8th June 2020).

[18] Marr, B., 2017. The Amazing Ways Google Uses Deep Learning AI. https://www.forbes.com/sites/bernardmarr/2017/08/08/the-amazing-ways-how-google-uses-deep-learning-ai/#389102923204 (accessed on June 8th, 2020).

[19] Faggella, D. Last updated: 2019. The State of AI Applications in Healthcare – An Overview of Trends. https://emerj.com/ai-sector-overviews/state-ai-applications-healthcare-overview-trends/ (accessed on 8th June 2020).

[20] Cireşan, D., Meier, U. and Schmidhuber, J., 2012. Multi-column Deep Neural Networks for Image. http://people.idsia.ch/~ciresan/data/cvpr2012.pdf (accessed on 9th June 2020).

[21] Operational Information for Commissioning, 2016. Diagnostic Imaging Dataset Statistical Release. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/08/Statistical-Release-27th-October-PDF-613KB.pdf (accessed on 9th June 2020).

[22] McKinney, S.M., Sieniek, M., Godbole, V. et al. International evaluation of an AI system for breast cancer screening. Nature 577, 89–94 (2020). https://doi.org/10.1038/s41586-019-1799-6 (accessed on 9th June 2020).

[23] Walsh, F., Last updated: January 2020. AI 'outperforms' doctors diagnosing breast cancer. https://www.bbc.co.uk/news/health-50857759 (accessed on 9th June 2020).

[24] Sanofi. 2019. The Ethics of AI in Healthcare. https://www.sanofi.com/en/about-us/our-stories/the-ethics-of-ai-in-healthcare (accessed on 17th July 2020).

[25] NHS, 2018. Breast Cancer Screening. https://www.nhs.uk/conditions/breast-cancer-screening/ (accessed on 9th June 2020).

[26] Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Three (accessed on 9th June 2020).

[27] Zhou LQ, Wang JY, Yu SY, et al. Artificial intelligence in medical imaging of the liver. World J Gastroenterol. 2019;25(6):672-682. doi:10.3748/wjg.v25.i6.672 https://pubmed.ncbi.nlm.nih.gov/30783371/ (accessed on 13th June 2020) and https://www.wjgnet.com/1007-9327/full/v25/i6/672.htm (accessed on 13th June 2020).

[28] Kanber, B., Nachev, P., Barkhof, F. et al. High-dimensional detection of imaging response to treatment in multiple sclerosis. npj Digit. Med. 2, 49 (2019). https://doi.org/10.1038/s41746-019-0127-8 . https://www.nature.com/articles/s41746-019-0127-8 (accessed on 13th June 2020).

[29] Siegersma, K., Leiner, T., Chew, D., Appelman, Y., Hofstar, L. and Verjans, J. 2019. Artificial intelligence in cardiovascular imaging: state of the art and implications for the imaging cardiologist. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6712136/ (accessed on 13th June 2020).

[30] Harris, S. 2018. Investment in Medical Imaging AI Companies Tops $500 Million. https://www.signifyresearch.net/medical-imaging/investment-in-medical-imaging-ai-companies-tops-500-million/ (accessed on 14th June 2020).

[31] Makary, M., Daniel., M., 2016. Medical error—the third leading cause of death in the US. https://doi.org/10.1136/bmj.i2139 https://www.bmj.com/content/353/bmj.i2139.full (accessed on 15th June 2020).

[32] A Comprehensive Look at Medical Malpractice Statistics. https://www.rosenbaumfirm.com/medical-malpractice-statistics.html (accessed on 15th June 2020).

[33] Clinical Decision Support Software Draft Guidance for Industry and Food and Drug Administration Staff. https://www.fda.gov/media/109618/download (accessed on 15th June 2020).

[34] Huss, M., 2018. Challenges of implementing AI in healthcare. https://peltarion.com/blog/data-science/challenges-of-implementing-ai-in-healthcare (accessed on 15th June 2020).

[35] Bender, E., 2019. Unpacking the Black Box in Artificial Intelligence for Medicine. https://undark.org/2019/12/04/black-box-artificial-intelligence/ (accessed on 16th June 2020).

[36] Watson David S, Krutzinna Jenny, Bruce Ian N, Griffiths Christopher EM, McInnes Iain B, Barnes Michael R et al. Clinical applications of machine learning algorithms: beyond the black box BMJ 2019; 364 :l886 https://www.bmj.com/content/364/bmj.l886.long (accessed on 16th June 2020).

[37] Liu, X., Faes, L., Kale, A., et Al. 2019. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. https://www.sciencedirect.com/science/article/pii/S2589750019301232 (accessed on 16th June 2020).

[38] Carfagno, Jack., 2019. IDx-DR, the First FDA-Approved AI System, is Growing Rapidly. https://www.docwirenews.com/docwire-pick/future-of-medicine-picks/idx-dr-the-first-fda-approved-ai-system-is-growing-rapidly/ (accessed on 16th June 2020).

[39] 2019. Statement from FDA Commissioner Scott Gottlieb, M.D. on steps toward a new, tailored review framework for artificial intelligence-based medical devices. https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-steps-toward-new-tailored-review-framework-artificial (accessed on 16th June 2020).

[40] Arevalo-Rodriguez, I., et Al. 2020. FALSE-NEGATIVE RESULTS OF INITIAL RT-PCR ASSAYS FOR COVID-19: A SYSTEMATIC REVIEW. https://www.medrxiv.org/content/10.1101/2020.04.16.20066787v1 (accessed on 17th June 2020).

[41] Watson, J., Whiting, P., Brush, J., 2020. Interpreting a covid-19 test result. BMJ 2020;369:m1808. doi: https://doi.org/10.1136/bmj.m1808 . https://www.bmj.com/content/369/bmj.m1808 (accessed on 17th June 2020).

[42] Stewart, S. 2019. Amazon pushes further into healthcare with NHS Alexa deal. https://www.thedrum.com/news/2019/07/10/amazon-pushes-further-healthcare-with-nhs-alexa-deal (accessed on 17th June 2020).

[43] Perez, S. 2019. Amazon Alexa launches its first HIPAA-compliant medical skills. https://techcrunch.com/2019/04/04/amazon-alexa-launches-its-first-hipaa-compliant-medical-skills/ (accessed on 17th June 2020).

[44] Youtube. 2019. The Age of A.I.. https://www.youtube.com/watch?v=UwsrzCVZAb8&list=PLjq6DwYksrzz_fsWIpPcf6V7p2RNAneKc (accessed on 7th May 2020).

[45] British Heart Foundation. 2019. Atrial fibrillation. https://www.bhf.org.uk/informationsupport/conditions/atrial-fibrillation (accessed on 18th June 2020).

[46] Arrhythmia Alliance. Information & Advice For Arrhythmia Patients. https://www.heartrhythmalliance.org/aa/uk/atrial-fibrillation (accessed on 18th June 2020).

[47] Turakhia MP, Shafrin J, Bognar K, et al. Estimated prevalence of undiagnosed atrial fibrillation in the United States. PLoS One. 2018;13(4):e0195088. Published 2018 Apr 12. doi:10.1371/journal.pone.0195088 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5896911/ (accessed on 18th June 2020).

[48] Stroke Association. 2017. Atrial fibrillation (AF) and stroke. https://www.stroke.org.uk/sites/default/files/atrial_fibrillation_and_stroke.pdf (accessed on 18th June 2020).

[49] Turakhia, M., Shafrin, J., Bognar, K., et Al. 2015. Economic Burden of Undiagnosed Nonvalvular Atrial Fibrillation in the United States. https://www.sciencedirect.com/science/article/abs/pii/S0002914915014290 (accessed on 18th June 2020).

[50] Hart, G., Pearce, L., and Aguilar, M. 2007. Meta-analysis: Antithrombotic Therapy to Prevent Stroke in Patients Who Have Nonvalvular Atrial Fibrillation. doi: 10.7326/0003-4819-146-12-200706190-00007. https://pubmed.ncbi.nlm.nih.gov/17577005/ (accessed on 18th June 2020).

[51] Tidy, C., Bonsall, A. 2017. Atrial Fibrillation. https://patient.info/heart-health/atrial-fibrillation-leaflet (accessed on 18th June 2020).

[52] Chandrasekhar, S. 2019. Apple Watch Sales History from 2015 to 2018: 3+ Years and 50+ Million Units in the Making. https://1reddrop.com/2019/01/07/apple-watch-sales-history-from-2015-to-2018-3-years-and-50-million-units-in-the-making/ (accessed on 18th June 2020).

[53] Canalys. 2020. Canalys: Worldwide smartwatch shipments grew 12% in Q1 2020 despite coronavirus. https://www.canalys.com/newsroom/canalys-worldwide-smartwatch-shipments-q1-2020 (accessed on 18th June 2020).

[54] El-Amrawy, F., and Nounou, M. 2015. Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial?. Doi: https://doi.org/10.4258/hir.2015.21.4.315 https://synapse.koreamed.org/DOIx.php?id=10.4258/hir.2015.21.4.315 (accessed on 18th June 2020).

[55] Hannon N, Sheehan O, Kelly L, et al. Stroke associated with atrial fibrillation--incidence and early outcomes in the north Dublin population stroke study. Cerebrovasc Dis. 2010;29(1):43-49. doi:10.1159/000255973 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2914401/ (accessed on 18th June 2020).

[56] Tison GH, Sanchez JM, Ballinger B, et al. Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch. JAMA Cardiol. 2018;3(5):409–416. doi:10.1001/jamacardio.2018.0136 https://jamanetwork.com/journals/jamacardiology/article-abstract/2675364?redirect=true (accessed on 18th June 2020).

[57] Singh, A. 2018. Detecting Atrial Fibrillation with the Apple Watch: Our Clinically Validated Results. https://blog.cardiogr.am/detecting-atrial-fibrillation-with-the-apple-watch-our-clinically-validated-results-ea66163e0fa6 (accessed on 18th June 2020).

[58] Apple. 2020. Heart health notifications on your Apple Watch. https://support.apple.com/en-us/HT208931 (accessed on 18th June 2020).

[59] Fox, M. The new Apple Watch has a heart monitor and the FDA approves. https://www.nbcnews.com/health/health-news/new-apple-watch-has-heart-monitor-fda-approves-n908976 (accessed on 18th June 2020).

[60] Capritto, A. 2019. The FDA just cleared an iPhone ECG sensor that beats the Apple Watch. https://www.cnet.com/health/the-fda-just-cleared-an-iphone-ecg-sensor-that-beats-the-apple-watch/ (accessed on 18th June 2020).

[61] Loon, I. 2018. FibriCheck receives FDA clearance for its digital heart rhythm monitor. https://www.fibricheck.com/fibricheck-receives-fda-clearance-for-its-digital-heart-rhythm-monitor/ (accessed on 18th June 2020).

[62] Gartner. 2020. Gartner Says Global Smartphone Sales Fell Slightly in the Fourth Quarter of 2019. https://www.gartner.com/en/newsroom/press-releases/2020-03-03-gartner-says-global-smartphone-sales-fell-slightly-in (accessed on 18th June 2020).

[63] Turakhia, M., Desai, M., Hedlin, H., et Al. 2019. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. doi: https://doi.org/10.1016/j.ahj.2018.09.002 https://www.sciencedirect.com/science/article/pii/S0002870318302710 (accessed on 18th June 2020).

[64] American College of Cardiology. Apple Heart Study Identifies AFib in Small Group of Apple Watch Wearers. https://www.acc.org/latest-in-cardiology/articles/2019/03/08/15/32/sat-9am-apple-heart-study-acc-2019 (accessed on 2nd July 2020).

[65] https://www.populationpyramid.net/united-states-of-america/2009/ (accessed on 6th July 2020).

[66] Stanford University. 2019. Apple Heart Study demonstrates ability of wearable technology to detect atrial fibrillation. http://med.stanford.edu/news/all-news/2019/03/apple-heart-study-demonstrates-ability-of-wearable-technology.html (accessed on 6th July 2020).

[67] American Heart Association. Apple Heart Study Clinical Trial Details. https://professional.heart.org/professional/ScienceNews/UCM_503785_Apple-Heart-Clinical-Trial-

Details.jsp (accessed on 6th July 2020).
[68] World Population. https://countrymeters.info/en/World (accessed on 6th July 2020).

[69] United States of America (USA) Population. https://countrymeters.info/en/United_States_of_America_(USA)#:~:text=As%20of%201%20January%202020%2C%20the%20population%20of,compared%20to%20population%20of%20330%2C377%2C563%20the%20year%20before. (accessed on 6th July 2020).

[70] O'Dea, S. 2020. Number of smartphone users worldwide from 2016 to 2021. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (accessed on 6th July 2020).

[71] The National Institute of Neurological Disorders and Stroke. 2019. Atrial Fibrillation and Stroke Information Page. https://www.ninds.nih.gov/Disorders/All-Disorders/Atrial-Fibrillation-and-Stroke-Information-Page#:~:text=The%20risk%20increases%20with%20age.%20In%20people%20over,have%20AF%20by%20approximately%20one-half%20to%20two-%20thirds. (accessed on 6th July 2020).

[72] Stroke Belt. Stroke Survival Rate. https://strokebelt.org/stroke-survival-rate/ (accessed on 6th July 2020).

[73] Brooks, M. 2018. FDA OKs AI Device to Detect Diabetic Retinopathy. https://www.webmd.com/diabetes/news/20180412/fda-oks-ai-device-to-detece-diabetic-retinopathy#:~:text=Diabetic%20retinopathy%20causes%20vision%20problems%20in%20people%20with,care%2C%20the%20FDA%20said%20in%20a%20news%20release. (accessed on 10th July 2020).

[74] Abràmoff, M.D., Lavin, P.T., Birch, M. et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. npj Digital Med 1, 39 (2018). https://doi.org/10.1038/s41746-018-0040-6 https://www.nature.com/articles/s41746-018-0040-6 (accessed on 10th July 2020)

[75] Frick KD, Gower EW, Kempen JH, Wolff JL. Economic Impact of Visual Impairment and Blindness in the United States. Arch Ophthalmol. 2007;125(4):544–550. doi:10.1001/archopht.125.4.544 https://jamanetwork.com/journals/jamaophthalmology/fullarticle/419213#:~:text=Economic%20Impact%20of%20Visual%20Impairment%20and%20Blindness%20in,without%20visual%20impairment.%20These%20differences%20remain%20highly%20significant. (accessed on 10th July 2020).

[76] Dastin, J. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G (accessed on 13th July 2020).

[77] Williams, D and Lavizzo-Mourey, R. 2016. Being Black Is Bad for Your Health. https://www.usnews.com/opinion/blogs/policy-dose/articles/2016-04-14/theres-a-huge-health-equity-gap-between-whites-and-minorities (accessed on 13th July 2020).

[78] Booth, R and Barr, C. 2020. Black people four times more likely to die from Covid-19, ONS finds. https://www.theguardian.com/world/2020/may/07/black-people-four-times-more-likely-to-die-from-covid-19-ons-finds (accessed on 13th July 2020).

[79] Egede, L. 2006. Race, Ethnicity, Culture, and Disparities in Health care. doi: 10.1111/j.1525-1497.2006.0512.x https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924616/ (accessed on 13th July 2020).

[80] Oxford English Dictionary. https://www.lexico.com/en/definition/medicine (accessed on 13th July 2020).

[81] Daniel Shiffman .2016. 9: Genetic Algorithms - The Nature of Code. 9: Genetic Algorithms - The Nature of Code. https://www.youtube.com/watch?v=9zfeTw-uFCw&list=PLRqwX-V7Uu6bJM3VgzjNV5YxVxUwzALHV (accessed on 14th July 2020)

[82] Piscasso, G. ML reinforcement learning 3 Types of Machine Learning: Supervised, Unsupervised and Reinforcement Learning. https://www.sailthru.com/marketing-blog/reinforcement-learning-machine-learning/ (accessed on 14th July 2020).

[83] Singh, P. AI capabilities in Image Recognition. https://towardsdatascience.com/ai-capabilities-in-image-recognition-7d79aec5222f (accessed on 14th July 2020).

[84] Vanian, J. 2020. Buzzy research lab OpenAI debuts first product as it tries to live up to the hype. https://fortune.com/2020/06/11/openai-artificial-intelligence-commercial-product/ (accessed on 14th July 2020).

[85] Panch, T., Mattie, H. & Celi, L.A. The "inconvenient truth" about AI in healthcare. npj Digit. Med. 2, 77 (2019). https://doi.org/10.1038/s41746-019-0155-4 https://www.nature.com/articles/s41746-019-0155-4#ethics (accessed on 14th July 2020).

[86] Thompson, C. 2019. Coders: Who They Are, What They Think and How They Are Changing Our World. ISBN: 978-1-52900-1900-1 (read throughout July 2020).

[87] Apple. 2018. ECG app and irregular heart rhythm notification available today on Apple Watch. https://www.apple.com/newsroom/2018/12/ecg-app-and-irregular-heart-rhythm-notification-available-today-on-apple-watch/ (accessed on 14th July 2020).

[88] Husten, L. 2019. Beware the hype over the Apple Watch heart app. The device could do more harm than good. https://www.statnews.com/2019/03/15/apple-watch-atrial-fibrillation/ (accessed on 14th July 2020).

[89] DeepMind. 2020. AlphaGo - The Movie | Full Documentary. https://www.youtube.com/watch?v=WXuK6gekU1Y (accessed 20th May 2020).

[90] Silver, D., Hubert, T., Schrittwieser, J., et al. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 07 Dec 2018: Vol. 362, Issue 6419, pp. 1140-1144 DOI: 10.1126/science.aar6404 https://science.sciencemag.org/content/362/6419/1140 (accessed on 14th July 2020).

[91] DW Documentary. 2019. Artificial intelligence and its ethics | DW Documentary. https://www.youtube.com/watch?v=Izd2qOgOGQI (accessed on 14th July 2020).

[92] Bailey, B. Aiken Standard (Newspaper) - May 1, 2014, Aiken, South Carolina. Accessed via newspaperarchive.com. Tech companies pour resources into artificial intelligence. https://newspaperarchive.com/aiken-standard-may-01-2014-p-51/ (accessed on 14th July 2020).

[93] Dankwa-Mullan I, Rivo M, Sepulveda M, Park Y, Snowdon J, Rhee K. Transforming Diabetes Care Through Artificial Intelligence: The Future Is Here. Popul Health Manag. 2019;22(3):229-242. doi:10.1089/pop.2018.0129 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6555175/ (accessed on 14th July 2020).

[94] Freethink. 2018. Saving Lives with AI | Freethink. https://www.youtube.com/watch?v=VePHPymCy2U (accessed on 14th July 2020).

[95] The Size and Quality of a Data Set. https://developers.google.com/machine-learning/data-prep/construct/collect/data-size-quality (accessed on 14th July 2020).

[96] The Economist. 2019. Is this the future of health? | The Economist. https://www.youtube.com/watch?v=jZg5QhL3Ckc (accessed on 14th July 2020).

[97] Bloomberg QuickTake Originals. 2018. The Rise of AI. https://www.youtube.com/watch?v=Dk7h22mRYHQ (accessed on 14th July 2020).

[98] World Medical Innovation Forum. 2019. 2019 WMIF | Last Mile: Fully Implementing AI in Healthcare. https://www.youtube.com/watch?v=hmcvwjHHnTA (accessed on 14th July 2020).

[99] Seto, E., Leonard, K., Cafazzo, J. 2012. Developing healthcare rule-based expert systems: Case study of a heart failure telemonitoring system. DOI: https://doi.org/10.1016/j.ijmedinf.2012.03.001 https://www.sciencedirect.com/science/article/abs/pii/S1386505612000561 (accessed on 14th July 2020).

[100] Harvard Medical School. 2018. MD vs Machine: How Artificial Intelligence Will Transform Medicine. https://www.youtube.com/watch?v=nUReWwR4FqM (accessed on 14th July 2020).

[101] Havard Medical School. 2019. MD vs. Machine: Artificial intelligence in health care. https://www.youtube.com/watch?v=xSDfma4VEx8 (accessed on 14th July 2020).

[102] Matt Lungren MD. 2018. https://www.youtube.com/watch?v=Gigd1rkZTSE (accessed on 14th July 2020).

[103] RE•WORK. Deep Learning in Medical Imaging - Ben Glocker, Imperial College London. https://www.youtube.com/watch?v=2_Jv11VpOF4 (accessed on 14th July 2020).

[104] Google Developers. Bringing AI and machine learning innovations to healthcare (Google I/O '18). https://www.youtube.com/watch?v=JzB7yS9t1YE (accessed on 14th July 2020).

[105] 2020. AI in Healthcare. http://aipodcast.education/ai-in-healthcare (accessed on 14th July 2020).

[106] Artificial Intelligence – Overview. https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm (accessed on 14th July 2020).

[107] Castrounis, A. Machine Learning: An In-Depth Guide - Overview, Goals, Learning Types, and Algorithms. https://www.innoarchitech.com/blog/machine-learning-an-in-depth-non-technical-guide (accessed on 14th July 2020).

Source Evaluation / Literature Review

My source evaluation / literature review will consist of the source number, a summary of what the source says, and a discussion on its reliability. Relevant links between sources and between sources and the project will be discussed where appropriate. As there is no direct research that determines the overall current socio-economic impact of artificial intelligence in medicine, there is little literature about the subject that currently exists. However, there is a lot of literature covering various facets of AI in medicine such as AI in mammography, primarily research into developing AI systems for a particular field. As such, the sources used in the dissertation serve as a literature review as they encompass relevant research into the field. As such my source evaluation will serve as an evaluation and a literature review. However, it is important to note that this is not considered a part of the dissertation but rather as a part of the wider process involved with the extended project qualification.

[1] The piece takes multiple approaches to estimating the economic value of artificial intelligence in the next ten years. It cites identifying new drugs as one of its future uses. The piece is written by a team of reputable authors who are qualified in economics and the methods used to estimate the economic value are sound. However, the piece is based on speculation alone and the piece explains that "calculating a single accurate estimate of AI's economic effect is difficult (if not impossible).". Prediction of the future is unreliable but in the context of this dissertation, the estimate conveys the idea that AI is predicted to be worth a large amount in the near future. The piece was released in 2016 and as such is slightly outdated, but the estimate is still valid to reference.

[2] This report by PricewaterhouseCoopers is written by a team of highly qualified economists and is a comprehensive piece that models the impact of AI to estimate its value in the future. The impact is assessed thoroughly analysing a wealth of data and creating a strong model. As such, this is perhaps the most accurate model one can find that assesses future worth. However, it is inherently limited by the fact that it is trying to predict the future and as such will probably be proven incorrect given time as predictions are very rarely precisely correct. However, I have utilised the source as I have clearly stated that it is an estimate of the future and as an estimate, this piece is perhaps the most thorough (and thus probably best) estimate one can obtain.

[3] The piece discusses the current state of artificial intelligence in various sectors such as healthcare and goes on to analyse possible future scenarios involving AI. In a survey conducted amongst business leaders, the report found that AI was most advanced in use in the healthcare sector (although it was not significantly more advanced than other industries), although they still concluded that AI was still far from being widely used in the industry and that its use was exploratory. The survey is trustworthy as it spans 203 executives and thus will help paint a decently accurate picture of current AI adoption levels. However, the small sample size means that this view of the healthcare sector is a rough one. As the survey does not focus particularly on healthcare, the picture may be slightly off, but it serves as an apt comparison between AI adoption in medicine and other industries. The paper comes from The Economist, a highly reputable economics source and thus the purpose appears to be merely to analyse AI adoption levels, making it a reliable resource.

[4] This article details some of the ethical considerations of artificial intelligence, such as the potential of furthering of racial bias, enabling immoral deeds such as use of AI powered military drones and potential to cause high levels of unemployment. Particularly relevant to this dissertation are the points it makes about poor training data reinforcing human bias and the need to establish precedent for the legal responsibilities of an AI system. This piece is however not a serious academic discussion and more of a voicing of fears that may become realised in the future. Whilst the focus is primarily on the future, it does highlight some instances of recorded unethical AI such as when a piece of software determined that a black person was more likely to commit a crime. The piece uses a few of these use cases to back up its argument lending some credibility, but the focus of the piece is mainly on future concerns, which are not highly relevant to my project. The piece's purpose may be to get people to read the article, possibly leaving room for exaggeration, as well as to start a discussion. The piece is not extremely reliable as a lot of the content is opinionated and coming from somebody who is not an AI expert, but the piece was included as the fears voiced are common and

consistent with others. Bossmann's arguments are justified with cases and is of some relevance to social drawbacks of AI in healthcare.

[5] The definition of artificial intelligence is obtained from a highly reputable dictionary and is aligns with my own knowledge and other definitions.

[6] The definition of data analysis is obtained from a highly reputable dictionary and is aligns with my own knowledge and other definitions. Understanding of data analysis is useful when analysing models such as neural networks which are used in healthcare.

[7] Jurafsky and Martin describe, in detail, the structure of a neural network and the use of an optimisation technique known as gradient descent to ensure that the network outputs near optimal values. They then go on to describe how this can be used to in language models that predict which word to write next given some number of previous words. This paper is a detailed academic piece published by Stanford University. The piece is extremely well sourced and everything it says is consistent with my own knowledge (which I partially acquired from programming my own neural network to develop the tool for [Figure 4]) as well as a wide range of other resources. The purpose of this piece is clearly to inform and teach, and links in with widely known knowledge of the field of machine learning. It demonstrates one use of AI in natural language processing. This source is, in short, very reliable.

[8] This YouTube video series describes the basic details of artificial neural networks and does so well, explaining the mathematics that underly the process of some machine learning models that make use of a neural network. Sanderson is an undergraduate in mathematics from Stanford University and is known to produce a wide range of reliable videos on computer science and mathematics. Everything appears to be well researched including an interview with a doctor whose PhD is in deep learning. Once more, the piece is consistent with widely accepted knowledge in the field and also aligns with everything source [7] details. The purpose is clearly educational and as such is a reliable source.

[9] Hallström details how a feedforward neural network functions, explaining its architecture and the concept of weights and biases. This resource has been very useful in terms of gaining a deeper understanding of how machines learn and thus their limitations and accuracy which has helped me in identifying weaknesses of AI in healthcare. The piece also goes into detail about the gradient descent algorithm, an optimisation technique for neural networks. This piece is consistent with general understanding of the topic and my own knowledge. The article is published on medium, a website that is a reputable resource for machine learning and data science articles.

[10] Machado describes the use of supervised, unsupervised and reinforcement learning techniques in machine learning. The piece is not highly detailed but mentions the broad idea of each one and does so accurately as his supervised section matches sources such as [8]. From other research and my own knowledge, the existence of unsupervised and reinforcement methods is well documented with algorithms such as k-means clustering and reinforcement learning and their use cases such as chess algorithms being well documented. As this piece is only used to show that other machine learning techniques exist, its inclusion is appropriate. However, unsupervised learning techniques are not common in medicine although still used as they are good at grouping similar data points for example and reinforcement techniques are practically non-existent in medicine.

[11] The article describes the existence of rule-based expert systems which can be thought of as systems that have had knowledge encoded into them by humans and are static models that can only be changed through human alteration of code. Such systems are well documented and do exist, with

one of the first AI computer systems to be developed for the medical domain being a rule-based system (called MYCIN). The piece goes on to compare rule-based systems to learning systems and points out that rule-based systems are advantageous as they do not suffer from the lack of transparency self-learning systems do (known as the "black box" problem). The piece points out that rule-based systems are flawed in some respects as they are difficult to maintain and rely on the information encoded into them being accurate. The information of this article seems to be accurate as it cites multiple sources, such as an image recognition algorithm that was fed poor training data and was unable to perform its task well. In short, all of its points are backed up and valid and thus can be viewed as a fairly reliable source. The idea of black box AI being a serious flaw is something that links into the inherent lack of trust physicians may place in a systems or conversely the excessive amount of trust a doctor is asked to place in the computer, even if the computer may be incapable of coping with data outside of the scope of its training sample. This idea of showing what areas an AI is looking at represents a possible way to increase trust in imaging AI.

[12] This a survey carried out across a range of companies and focuses on industrial artificial intelligence adoption. The 2018 study concludes that AI adoption is increasing rapidly in business but that a large amount of companies have little to no AI use. The study highlights that in some areas, AI is already used frequently but that in others, it is deployed in one or fewer aspects of the busines. The survey is reputable as it is published by a global institute that is known for studying specific markets and trends, as such the data is bound to be accurate and obtained from a range of companies. I did not include this source in my dissertation as it focuses more on the wider use of Artificial Intelligence as opposed to strictly in the healthcare sector, however it was utilised as part of my formative reading and helped me to understand the general levels of AI usage.

[13] This 2019 survey is essentially the same as resource [12] it just details the survey performed in a different year. It states that commercial AI usage has risen and that it shows benefits in industries that do use it. However, it also highlights that many have not made much use of the technology as barriers remain before use can be widespread. The survey is useful as it provides a broad overview of the state of commercial AI but has is not directly mentioned in this dissertation as the focus is on use in the wider business world, not healthcare.

[14] This paper by Jiang et al. was a highly useful source in understanding the current state of AI in medicine. The piece details current use cases and research into AI in healthcare. The piece is extremely thorough, with 69 citations and the pieces' explanation of techniques such as suppport vector machines, neural networks and deep learning align with other sources and general knowledge. It is published in one the most reputable journals, The British Medical Journal, and is peer reviewed which means that a high degree of trust can be placed in both the accuracy and reliability of this resource. Perhaps the only limitation is that the source was published in 2017, as AI in healthcare is rapidly evolving. However, 2017 is still relatively recent and the content broadly aligns with my impression of AI in healthcare today. This idea of prevalence of AI in both implementation and research for the healthcare industry links to the socio-economic impact of AI as the advancement and relative amount of research and implemented systems in the industry is linked to its overall impact.

[15] This journal piece describes the use of support vector machines (an AI technique) to categorise pathological gaits, particularly amongst stroke sufferers. The system achieved a 90.5% accuracy, showing promise for future advancements in the field. This study is performed using a rigorous and scientifically sound method that has a high accuracy level. It is also published in a reputable, peer-reviewed journal meaning that the information will have been checked by other experts which I concluded enables me to have a high degree of confidence in the results of this piece. The purpose is

to objectively present their research although there is always a small chance that it may be cherry-picked to suit their narrative (although unlikely as this piece appears to be exclusively academic). However, I have used the article as it is valid in the context of an AI system that accurately detects gait class.

[16] This news articles published in Forbes details the first FDA approved machine learning software for use in medicine, an artificial feedforward neural network produced by Arterys's to analyse medical scans of the heart. The piece interviews the founder of the platform. All in all, the piece is highly relevant as it describes the first FDA approval which provides context as to just how recent implementation of AI in healthcare system. The article may well exaggerate its capabilities as the focus is on interviewing the founder, somebody who is unlikely to be completely neutral in the matter, even if his arguments are primarily backed up by facts. The article importantly highlights that in order to gain FDA approval, it had to pass a range of rigorous tests to prove that it functioned just as effectively as humans, which shows how difficult it is to gain approval. This knowledge is highly useful as it links in to the socio-economic impact of research as if an AI system is able to perform at a similar level to a human, it is theoretically able to be integrated into the healthcare system subject to further research, but it at least shows that implementation is feasible for such algorithms. I used the article as it is reliable in the sense that it accurately shows the date of the first approval of a machine learning algorithm by the FDA and as it demonstrates the standard required for such a system to be approved.

[17] This chart from Statista shows that almost 50% of AI in marketing is used for data analysis. This resource was included because the focus of the project was originally going to be on healthcare and data analysis, however I concluded that focusing on solely healthcare was more appropriate. Statista is a somewhat reputable website for statistics, however it is a far less reliable source for non-paying members (such as myself) as it limits the access to the number of articles one may view and does not show which sources were used unless one pays for the premium membership option. That being said, I know that a large amount of AI in marketing is data analysis and as such this source is probably reliable, but it is difficult to verify this as I am unable to see the sources this chart uses. Therefore, I conclude that the information is probably reliable and is definitely correct in terms of AI in marketing involving data analysis, but that it is too unreliable to use. Furthermore, the information is largely irrelevant to my dissertation although the inference that AI is good at analysing large amounts of data (particularly about customers) is applicable to healthcare, particularly as other resources ([94] and [98]) say that AI is used to analyse the data of patients and that hospitals collect enormous amounts of data on patients.

[18] This news article by Marr details the training of a deep learning algorithm by Google that uses of 10 million unlabelled images to successfully identify cats. This shows the potential of AI to meaningfully analyse large amounts of data, something which is invaluable to the medical industry where many healthcare providers complain of an excess of data. The article is written by Marr, somebody who frequently writes about AI and is published in Forbes, a well-known news site which enforces journalistic due diligence. The article can be externally verified as Google have published a paper on this AI and one can also verify the article's claims about Google's current other uses of AI such as self-driving cars. Initially, my dissertation was going to have a focus on data analysis and healthcare, but I decided that the analysis would be of greater quality if I focused on healthcare. As such this article was not included as it lacked relevance as it does not talk much healthcare (although it does mention that DeepMind (a Google company) is working to develop imaging AI).

[19] This article is based on and includes a discussion between two machine learning and healthcare experts. The consensus is that AI is likely to become widespread in healthcare in the immediate

future but that its current implementation is far more limited. This is due to issues such as the black-box nature of AI but also due to the structure of the healthcare industry. They say that there is a lack of incentive to invest in research in the field because stakeholders have not seen enough to be convinced that the research will return their investment. This speaks to the problem of not enough research being translated into actual implementation yet, creating a repeating loop where there is a lack of proof to show that investment into AI is worthwhile, thus leading to lack of investment into research which leads to a lack of proof that AI is worthwhile. However, I disagree with the experts slightly here as I have seen a wide range of research into AI that shows extremely successful software. But I do generally agree and accept their point as they work in the industry although I feel that it isn't a lack of research that is the problem but more lack the lack of research being implemented. What they point out, which is invaluable, is that because it is such an emerging market, many investors are hesitant to invest in the field which leads to slow growth. However, this is an important point as it links to an underlying theme of the dissertation (that research does present socio-economic value as it leads to more investment from stakeholders and thus adoption of AI in healthcare). All in all, this resource is highly reliable as it is written by an expert who has interviewed the people who are most likely to understand in detail the current state of AI in healthcare.

[20] This academic piece details the creation of a deep convolutional neural network to attempt to recognise handwritten digits using the MNIST dataset, a set of 70000 handwritten digits. This network achieves an accuracy of 99.77% on the dataset, which is highly impressive, particularly when considering that the MNIST dataset is known to be imperfect (i.e. a very small number of digits can be misidentified by humans as they look very similar to other numbers). I have no doubt that the piece is accurate, as other state of the art networks are known to achieve similar performances. This paper is very thorough and describes both its architecture and training method in a way that gives me no reason to believe that this is anything but true. As such, the source is reliable and is relevant as it shows the potentially of AI to achieve near perfect accuracy in image recognition tasks. This is highly useful in healthcare as many doctors analyse images such as mammograms, X-ray scans or MRI scans to reach a diagnosis.

[21] This document release by the English National Health Service publishes a years' worth of scanning data from 2015-2016 and analyses a number of things such as time between a scan being performed and receiving a diagnosis, time between booking and receiving a scan and number of scans performed. This is a detailed official document and as such can be trusted to be reliable as the purpose of the document is to be accurate. If it were found that the document was altered deliberately, there would be a large governmental scandal and many people would lose their jobs. Simply put, the authors have no reason to lie and a very strong incentive (not to mention safeguards) to ensure that the piece is true. There is no doubt in my mind that this document is of the strongest validity and is highly useful as it, for example, highlights that some patients can wait up to twenty one days to receive the results of an MRI scan. This is something that AI could help to drastically speed up.

[22] This paper in the Nature journal talks about the development of an AI system that analyses mammograms in order to determine whether a patient has breast cancer. The system performed highly impressively, reducing false positives by 5.7% and reducing false negatives by 9.4% in the USA. The system also performed favourably when compared to experts in the UK. This piece was published in one of the most respected journals (Nature) which publishes peer-reviewed work. This means that the claims made in this article can be trusted and are accurate. This case is highly relevant to the dissertation as it shows that AI can outperform human doctors as well as taking

milliseconds to analyse a scan. Research such as this does have socio-economic implications as it demonstrates the fact that AI can free up doctors to perform other tasks where they are needed as well as reduce the overall cost of healthcare as once such a system is completed, maintenance costs are negligible when compared to the cost of a radiologist. Research of this nature should also help to address the hesitancy medicinal stakeholders feel in investing in the technology as they feel that it is unproven (as shown in [19]). This article was published in 2020 showing just how recent this breakthrough is, helping to understand why there is such limited current use of AI in the field. This research case is used as the typical example of the impact of AI research and therefore is a valuable resource to the overall dissertation.

[23] This BBC News report summarises the main points of the Nature article that is used at detecting whether breast cancer is present ([22]). This piece focuses much more on the socio-economic implications of AI and it describes concerns over divulging of sensitive health data to third parties as well as hinting at the fact that there may be problems when such a system is implemented. It also discusses the benefits of the system as it is tireless and is well suited to optimise the current system by replacing one of the doctors that normally analyse such a scan. In this instance, an expert believes that the AI will not lead to a lack of demand for radiologists as there is already a large shortage of them. However, I would argue that this is probably less applicable more broadly and imaging AI in general will lead to far fewer doctors being needed. However, in the very long term, better accuracy and cheaper medical bills can only be a good thing even if there are large social implications to reducing the demand of doctors (such as lack of advancement in medical research and large unemployment levels). This piece is published by the BBC, which is a trusted news organisation which has an agenda of impartiality. The article gives no opinions but instead interviews various experts who give their viewpoints on the potential socio-economic impact of this system. As such, this article is highly relevant to this dissertation.

[24] The ethics of AI and potentially social concerns regarding use of Ai in healthcare are discussed. For example, it highlights the potential for unrepresentative data to limit the effectiveness of AI. Facial recognition software that performs far worse on people of colour is one of the examples it uses to demonstrate that this is a flaw with AI that is being seen. This can be overcome with a greater data pool, but there are fears that this issue may never be fully combatted effectively. It also details the worry of the black box nature of an AI, something which other resources on the subject also highlight including [19]. The piece summarises the European Union’s guidance on ethical AI. Whilst the piece is reliable and accurate so far as justifying all of its points well and as these points are generally agreed on, the piece does have limitations. This is because Sanofi are creating their own ethics code on AI in healthcare which of course would place them in a good position to understand the implications of the technology, but it also gives them a reason to be unfairly negative about the ethics of AI and the EU’s guidelines in order to demonstrate the need for a better set. Therefore, I have not directly used the piece as I feel that this is a serious bias. However, I do feel that the source is valid as the ethical concerns it mentions are shared by other experts such as [19], I just feel that the resource might have an incentive to not be neutral. I have used it as formative reading and have considered the ethical drawbacks it cites, all of which are also mentioned by other sources.

[25] This piece by the National Health Service describes the standard protocol of breast cancer screenings in the UK, such as that women are typically screened once every three years and that one can wait up to (but typically no more than) two weeks to receive the results. As the piece is published by the NHS and aimed at patients, its purpose is clearly informative as it wishes to provide patients with accurate information which makes the source of high validity. It is also relevant as an

AI that speeds up mammogram analysis would allow for a shorter wait time for results but also the possibility of screening more frequently than once every three years which also makes this source relevant.

[26] This is a set of statistics published by the charity Cancer Research UK about the survival rates of breast cancer at each stage of diagnosis (e.g. survivability if it is identified at Stage 1 compared to Stage 3) as well as other survival statistics. The statistics highlight that the early diagnosis of breast cancer leads to much greater survival rates with 97.9% of those who were diagnosed at stage 1 surviving for more than 5 years, compared to 72% and 26.2% for stage 3 and stage 4, respectively. The source comes from one of the largest charities in the UK whose primary goal is research into cancer to improve survival rates. As such, statistics of this sort are highly likely to be accurate as the integrity of such a charity would be diminished if such statistics were false. Of course, one could argue that they may be biased to exaggerate the risk of death from cancer and the need to get checked up regularly but this is clearly no the case as the data stems from the office of national statistics. What this does show is the potential of AI that speeds up mammography to allow for more frequent scans for breast cancer which would mean that more cases were identified early, leading to a larger number of people surviving breast cancer.

[27] This peer reviewed paper details various types of machine learning techniques and phrases such as convolutional neural networks, deep learning etc. The piece then describes the potential use of these systems in medical imaging such as the diagnosis and detection of liver disease. The paper concludes by highlighting some hindrances to future implementation of AI in healthcare, such as the fact that AI may be good at performing certain tasks but that it lacks the higher levels of awareness and context and thus may be flawed when dealing with complications. The piece is written by a team of academics and peer-reviewed which means that it is safe to conclude that the content is accurate. Furthermore, the content aligns with other resources leading me to believe that this source is valid. It also links into this recurring idea of AI showing potential but not yet being implemented.

[28] This paper describes the use of a high dimensional neural network in treating multiple sclerosis which outperforms models of lower dimensionality. Once more this piece is published in a reputable journal by qualified authors and has been peer-reviewed which gives me a high level of confidence in the validity of this source. This idea of greater dimensionality (i.e. more complex structure) being effective indicates that neural networks can be highly effective at performing tasks the human brain does, given enough scope to carry out these tasks, showing the potential of systems in a wide range of advanced, yet limited in scope, tasks. This recurring use of AI in medical imaging shows potential social benefits as, for example, a general X-ray scan could theoretically be scanned by multiple AI systems in a matter of milliseconds in order to screen for a whole range of conditions that may have never been detected otherwise.

[29] This paper reviews a range of state-of-the-art cardiovascular imaging AI which have superhuman or near human performance. This paper concludes that AI holds incredible potential to advance the field but that its implementation has been limited due to a number of reasons such as approval being slow and a lack of research that shows the cost effectiveness of AI in medicine. The contents of this paper are very reliable as it summarises a range of reliable sources expertly and as it is published in a respected journal. The paper is also current as it was published in 2019, meaning that the viewpoint of imaging AI will be broadly accurate right now.

[30] This news article talks about the levels of funding that healthcare AI start-ups were receiving, with the headline being that over half a billion dollars was invested in the first half of 2018 alone,

although there does appear to be less interest than in 2016. However, this level of funding is important is at it may negate one of the key criticisms of the healthcare AI mentioned in [29] and [19] about the lack of funding into research discouraging stakeholders from adopting the technology. The research this report is based on (which is performed by the same company as the one writing the article) is thorough and verifiable in regards such as its reference to founded companies and through public records of funding. As such, I conclude that the report is useful from the perspective of showing that there is a large level of investment into AI in healthcare right now and as it is published by a company who focus on giving insight into technology usage in healthcare.

[31] This piece is published in The British Medical Journal, a peer reviewed and highly respected medical journal, which concludes that medical error was the third leading cause of death in the USA in 2016. As a result of the reputation of both the journal and its academic rigour as well as thorough ness of the paper itself, I conclude that the source is valid and its conclusion reasonable. This piece is of course relevant to AI as AI systems that outperform doctors could arguably reduce medical error if programmed well or increase it through poor design decisions such as the mentioned lack of diversity in training data in [29].

[32] This piece is a collection of reliable resources that give an insight into the number of medical malpractice claims per year. All the sources it uses are reliable and show a very large number of annual cases although the source has a valid reason to show bias as it is published by a law firm that claim to "HOLD DOCTORS AND HOSPITALS ACCOUNTABLE" which would certainly give it a reason to be biased. However, the data it presents is verifiably accurate such as the fact that John Hopkin's University estimates that around 250,000 cases of medical malpractice occur annually. Therefore, I have included the source as these statistics give a reasonable image of the volume of malpractice claims.

[33] This is a draft of potential FDA guidance regarding the use of software in a clinical environment that shows a willingness for approval of software in "non-serious" situations although it shows an openness and uncertainty about software in more critical situations. It is important to note that this is a draft and not actual guidance and therefore can only be viewed as the general thought process of the FDA. The draft is current as it was released in September of 2019 which suggests that the FDA's current stance is similar and as such is a relevant document as the socio-economic impact of research into AI is directly linked to the ability of these research projects to be implemented by healthcare providers (which would require FDA approval in the USA). And of course, this document is bound to be valid as it is written by the FDA.

[34] Huss' blog post about the potential hurdles of implementation of AI in healthcare is a well thought out piece that points out a number of important flaws such as the greater need for transparency (something which is legally required of for AI healthcare devices in the EU). He also highlights the problem of doctors being unable to trust AI and its black box nature (which is a recurring theme across multiple resources). Despite the general content and critiques of AI being fair, the piece does not back many of its arguments up (although it does correctly highlight relevant legislation) and is somewhat opinionated. Despite the lack of formality that comes with a blog post though, the flaws of AI are agreed upon by other resources and experts such as [19] and [98]. Furthermore, the blog has the potential to be biased as the platform is developing a platform that addresses one of the issues described in the blog, giving them incentive to exaggerate the downsides of AI and particularly of the downside of "engineering debt", the issue the platform claims to combat. However, I have included the sources as the general critiques are valid and shared by other experts even though there is a conflict of interest.

[35] This article describes a common problem of AI in healthcare, its lack of transparency(i.e. the fact that doctors are unable to see what it is doing and how it reaches its decision) or “black box” nature. The article features interviews of professionals who sit on both sides of the fence. A computer scientist at the Massachusetts Institute of Technology calls the black box problem a “myth” something which I feel is simply untrue given the sources such as [19] and [98] which feature healthcare and machine learning leaders who all say that the black box problem is one the largest issues when it comes to implementation. Furthermore, legislation such as the one in the EU ([34]) show just how seriously the problem is viewed. However, the piece goes on to highlight critics who argue that if a doctor can not trust the result of an AI then neither can a patient. The piece also says that deep learning companies are actively in the processing of developing tools to interpret deep networks although very few tools are currently available. Given the suggestion that there appear to be a trade off between interpretability and accuracy, one scientist (Goldenberg) concludes that the best approach is to build the best possible system and then develop tools to make it as interpretable as possible, without compromising the accuracy of the network. This article is valid as it appear to be neutral (due to the fact that it features a range of viewpoints from experts) and is extremely useful as it provides insight into one of the biggest problems with AI in healthcare and the contrasting viewpoints of experts on the matter. And ultimately, I agree with the viewpoint of Goldenberg although I do accept the validity of the other viewpoints presented here as well. But to me it shows that the black box issue is a problem that has yet to be fully overcome but could quite possibly be combatted before too much longer.

[36] This peer reviewed British Medical Journal piece assesses one common issue with AI: its black box nature. It concludes that machine learning algorithm are inherently black box in nature and that for legal, social, and moral reasons, this must be altered. Once more, this article’s content is reputable due to the standard of the journal it is published in and the fact that it is peer-reviewed. Furthermore, this problem with the black box and the difficulty to trust a machine that does not justify its answers leading to potentially catastrophic results is echoed across the community as seen in other resources such as [35]. As such, the content of this piece is accurate and hence the source is valid for inclusion.

[37] This is one of the most substantial pieces of literature on research into machine learning in healthcare. This piece provides a comprehensive analysis of a range of 69 studies about AI systems in healthcare and concluded that the algorithms operated at a level similar to healthcare professionals. This lends massive strength to the underlying idea of this essay: that the ability of researched AI to outperform or perform comparably to doctors is typical for such projects. As this study analyses a range of other papers (all of which it links to), the authors have clearly performed the analysis they claim to have and therefore this is source is both valid and highly relevant to the essay.

[38] This news article describes how the use of the first FDA approved AI technology, IDx-DR, used for lesion detection is rapidly growing in its use. The article is verifiably accurate as papers about the accuracy of IDx-DR as well as articles about IDx-DR being an FDA approved healthcare AI are easy to find, showing that the author has done due diligence. Therefore, the accuracy can be trusted, even if the interview with the founder of the software may lead to a slightly optimistic view of its capabilities. However, the article’s claim of IDx-DR being the first FDA approved AI system is perhaps slightly misleading as Artery’s software appears to qualify as the first approved medicinal AI. However, the article does go on to clarify that IDx-DR is the first autonomous AI approval as it can be used without a qualified professional present (i.e. it is allowed to diagnose as opposed to just advise).

[39] This is a statement published by the FDA announcing their intentions to work with the community to produce a set of guidelines to ensure ethical and good practices are maintained when producing AI in healthcare. The statement also underlines the FDA's willingness to embrace the technology as they feel it has tremendous potential benefits. This statement is of great relevance as it shows that there is a lack of clear guidance available about medical AI, which will of course lead to far fewer implementations of AI, that ethical AI is of great importance (and that there is a need for guidelines to overcome some of the hurdles medical AI presents), and that the FDA is willing to accept and work with companies to implement AI in the field. Therefore, the source is of great relevance.

[40] This piece reviews existing work to attempt to calculate the false negative rate for a commonly used test for COVID-19. The piece concludes that there is an up to 29 percent false negative rate for the RT-PCR COIVD-19 test. However, this piece is highly unreliable as it is published very early on into the lifespan of the COVID-19 pandemic and is using data from other samples that will be limited by errors that are not the fault of the test such as human error and mixing up of samples. Furthermore, this work has not been peer reviewed and the publisher even stresses that this number should not viewed as a fact and that the evidence is severely limited for a range of reasons. As such, the source lacks validity and confidence in the results. However, it is of some use as it shows that tests, which are similar to an AI in their black box nature (i.e. they often return yes or no) are also flawed and are still widely used for obvious reasons. Admittedly, this comparison is not entirely fair as the process behind the test is understood and often the accuracy of the test is also understood (although the accuracy of the COVID-19 test is not completely agreed upon).

[41] The British Medical Journal piece talks about the accuracy of COVID-19 tests and how a negative or positive result should be interpreted. The piece references source [40] and concludes that the between 2% and 29% false negative rate may be an underestimate. It also highlights studies which show that other forms of the COVID-19 test have a greater false negative rate than 29%, which suggests that the tests are at least somewhat flawed. However, whilst this is published in a reputable peer-reviewed journal, the piece is unlikely to be completely accurate as it is based on preliminary and small studies and was written when COVID-19 was a very new phenomenon. However, it does at least show that the COVID-19 tests are not perfect and lends strength to [40]'s conclusion.

[42] This news article describes an agreement between Amazon and the NHS which will see its voice assistant device (known as "Alexa") relay health advice from the NHS website should a user request such information. This shows AI's push into care and the ability of natural language processing AI to make a real difference particularly to those who lack digital literacy. However, devices such as "Alexa" fall somewhat outside of the scope of the essay as they do not meet the definition of medicine although they do help to illustrate the use of AI and NLP in areas such as care. The article is accurate as the information has been publicly confirmed by all parties involved.

[43] This article describes how, following a clinical trial, Alexa has allowed certain developers to cerate medical skills for the Alexa app as long as the comply with relevant medical standards. Similar to [42], this shows the ability for NLP software to make a difference both in the fields of medicine and care. The contents of this article have been confirmed by Amazon, lending the source validity.

[44] This introductory video series on Artificial Intelligence, published and created by YouTube, covers a wide range of subjects including how AI functions as well as applications in healthcare. It shows an example of an artificial arm that uses electronic singles from a user and an AI to give the wearer more advanced control of their arm and hand than a typical replacement limb. This is a series designed to entertain and also inform and the series is clearly very well researched as it is published

by such a large company and clearly has a large budget. I did not include the source directly as the piece tends to provide a broad overview of each area and as the medicinal aspect was highly experimental (and less suitable to analyse than the breast cancer diagnosis AI as there were less resources available on the "Skywalker" hand). However, the source is still valid and helped me shape and enhance my understanding of the topics and themes in this dissertation.

[45] This piece by the British Heart Foundation is designed to inform about the condition of Atrial Fibrillation, one of the most common cardiovascular conditions. The British Heart Foundation describes symptoms of AF, types of AF (such as paroxysmal) as well as treatments. This piece can be trusted as it clearly has a purpose to inform, because the British Heart Foundation have a high degree of knowledge and expertise in cardiovascular diseases and because everything it says about AF can be verified by other reputable health organisations. Therefore, the piece is useful as it provides a detailed explanation of the condition and that it can be treated effectively and that the condition can be intermittently present.

[46] Similar to [45], this piece also talks in detail about many aspects of AF and provides advice for patients including about treatment methods. The Arrythmia Alliance's goals are to "to improve the diagnosis, treatment and quality of life for all those affected by arrhythmias" (such as AF) and as such are highly likely to be informative and fair, something which is backed up by the fact that their information matches that of [45].

[47] This paper estimates that 5.3 million people had AF in the USA in 2009, of which 0.7 million were undiagnosed (13.1%). The paper uses a strong and reliable methodology that can be replicated as the data it is analysing is publicly available. The piece is authored by a range of experts, including an academic form Stanford University's School of Medicine and therefore is trustworthy. As such, I have strong confidence in the accuracy of the results of this paper. This piece is very important as it highlights the great benefits of more successfully identifying AF.

[48] Guidance by Stroke.org talks about what AF is as well as its prevalence and treatment. It points out that AF sufferers are at a five times greater risking of having a stroke (something which can be reduced through treatment) and that an estimated 1.2 million British people suffer from AF. The piece is reliable as it is published by a charity who aim to raise awareness, reduce the number of strokes, and improve the quality of treatment for stroke victims. The purpose of the piece is also very clearly to inform which leads me to see the source as reliable. This piece is important as it shows the need for AF sufferers to seek treatment to reduce their stroke risk which makes the diagnosis of AF patients a very important matter.

[49] This piece analyses the healthcare costs of those with undiagnosed AF and compares it those with diagnosed AF and finds that "incremental cost burden of undiagnosed nonvalvular AF is $3.1 billion" (in the USA, when compared to those with diagnosed AF). The methodology of this paper is sound and as such I have a high degree of confidence in the findings. Furthermore, this will be an underestimate for the overall economic costs as undiagnosed AF leads to a greater stroke risk which has economic ramifications beyond medical expenditure such as inability to work and poorer quality of life resulting from strokes that induce disability. In short, this shows the massive economic potential of a device such as the Apple Watch as it could have a large economic impact and a large social impact as well.

[50] This article provides an analysis of 13 randomised trials on the use of certain drugs to reduce the risk of strokes amongst AF patients and concludes that use of Warfarin can reduce stroke risk by 60% for AF sufferers. This shows an enormous economic benefit in improved diagnosis and

screening for AF as proper treatment can clearly reduce stroke risk and hence increase both quality and length of life of sufferers of the condition. This piece was written by a member of the University of Texas' medical department and has a sound methodology that leads to a high level of confidence in its findings. Thus, I deem the source to be both valid and relevant.

[51] This article describes what AF is, including symptoms and types of AF. Particularly relevant is its description of paroxysmal AF, where symptoms of AF are sometimes present but abate for periods of time, as paroxysmal AF is very common amongst those who have not been diagnosed with AF. The piece is published in a reliable medical cite and has been written by a doctor and reviewed meaning that the content can be trusted. The information presented is also consistent with other reliable sources such as [45] and [46], lending more validity to it. The purpose appears to be solely to inform.

[52] This article reviews Apple's public statements, Apple's revenue documents and various expert analyses to give a rough overview of the total sales of the Apple Watch. The piece uses a range of reputable sources and insights from various experts to roughly track sales and shipments of the watch to successfully show that the Apple Watch has sold many millions of units and continues to perform well for the company. As a result of its use of reputable sources and direct quotes from the company, the information presented is clearly reliable and hence the source is valid.

[53] This is a detailed analysis of the performance of smart watches such as the Apple Watch in the first quarter of 2020. It is performed by the reputable tech analysis firm Canalys and as such is likely to be well researched as the reputation of the company is dependent on the accuracy of such reports. The piece shows that the Apple Watch is the dominant smart watch brand and its high shipment numbers show the large potential of AF detection software across such a large group of people.

[54] This academic paper measures the performance of wearables such as the Apple Watch and compares their accuracy in measuring metrics such as steps and heart rate to professional devices. The study found that the Apple Watch had a 99.9% accuracy in measuring heart rate when compared to a clinical device which was the highest accuracy of all devices tested and shows that the Watch acts at the same level as advanced pulse measuring equipment. The piece is published by two academic individuals and all the methodology seems to be thorough and fair, leading me to conclude that its findings are valid.

[55] This paper analyses the proportion of strokes occurring in North Dublin that are related to AF. The piece concludes that 31.2% of strokes were AF related and that a prior stroke was twice as likely in AF suffers than non-AF suffers. The piece is comprehensive and has a fairly large sample size although the potential for selection bias is there (although steps were taken to minimise this risk). The study does appear to be reliable as it is written by a group of academics and because both the methodology and data seem to be proper. However, there is the potential that the true figure is not 31.2% as this number is based on 568 stroke incidents. Despite this, the estimate is still reliable and demonstrates the fact that AF drastically increases likelihood of stroke (and strokes have socio-economic ramifications as they can lead to disability and death).

[56] This paper summarises the development of a deep neural network that detects atrial fibrillation using the Apple Watch which achieves a sensitivity of 98.0% and a specificity of 90.2%. This paper is widely accepted and maintains thorough methodology and impressive results, showing the accuracy of the model. Furthermore, it is not the only Apple Watch AI model that has achieved similar results

lending validity to its results. The high accuracy of diagnosis shows potential as it enables passive and continuous measurements across a large population.

[57] This blog by Singh summarises the results of paper [56] and provides some additional detail and context about the paper. In essence, it summarises [56] and makes the papers contents easier to divulge. Having read [56], the contents of this article are clearly accurate and a reflection of [56], making it valid.

[58] This piece on the support section of Apple's website talks about the irregular pulse notification on Apple Watches that can be enabled. The support piece highlights that the feature is not supported in all countries but is in some. This source is valid as it is designed to inform an Apple Watch user and comes from the company who would know the finer details about the watch.

[59] This NBC news article details the unveiling of a new Apple Watch in 2018 which has the ability to detect AF. The piece goes on to say that the FDA has approved use of this technology in the watch. This article is published by reputable news organisation NBC and can be varied by other sources such as [Figure 2], [Figure 3] and the video recording taken from the conference in which the watch was unveiled. As such, the piece is clearly accurate.

[60] This news article talks about the FDA's approval of a small sensor that is used alongside a phone application to measure heartrate and can be used to detect conditions such as atrial fibrillation. The report can be trusted as the information in it can be verified through other articles, FDA records and Alivecor's statements (the company who created the product). Therefore, the source is valid.

[61] This piece on Fibricheck's website details the fact that the FDA has approved their atrial fibrillation detection software that makes use of a smart phone's built in camera. This information is likely to be accurate as it is published as it is published on Fibricheck's website, a medical organisation who would be penalised for spreading misinformation about their products, although it could also lead to bias and a view of the product that is overly positive. However, the broad content of the article can be confirmed as the product is available and as its approval is logged by the FDA. This software has the potential to save many lives as it makes use of a smartphone which has an audience of many billions (a much larger size than even the Apple Watch). Although it does come with the downside that constant monitoring of pulse is not possible. It does link to the wider theme of AI making healthcare more accessible and available to a wider sample of people.

[62] The number of sales of smartphones in the last quarter of 2019 along with the overall sales in that year are analysed in this piece. Produced by Gartner, a reputable technology research organisation, this information is likely to be accurate as the purpose of the company is to publish accurate and reliable research (along with the fact that much of the data presented here can be verified through earning reports of each company). This piece shows the vast number of mobile phones sold each year, with around 1.5 billion sold in 2019. This is relevant as it shows that screening for conditions such as AF is available to a large demographic of people by making use of software such as to one mentioned in [61]. This is important as it enables doctors to screen a much larger pool of people for AF without having to distribute limited resources.

[63] This science direct journal piece describes the methodology of the Apple Heart Study, a study that makes use of the Apple Watch amongst a group of 400,000 people to scan for atrial fibrillation. It also describes the rationale and benefit behind such a study, highlighting the massive economic detriment of undiagnosed AF as well as the benefits of reduction in stroke risk. The contents are credible as the study was performed by reputable organisations showing that this method was implemented. However, the piece is limited as the study is payed for by Apple which is a conflict of

interest. Despite this, best academic practices were clearly adhered to. The large sample size is important to strengthen quality of data although the trail design is limited by the fact that it is not randomised and that the demographic may be skewed towards those who are young (and less likely to have undiagnosed AF).

[64] This article summarises the results of the Apple Heart study and includes an interview with one of the researchers involved in the trial. The contents of the article are accurate as the results have been widely reported. The piece highlights that only 57% percent of participants who received a notification sought healthcare, perhaps showing a lack of trust in the capabilities of the Watch and are certainly a flaw of the trial in the sense that it may have had a large impact on the accuracy of the watch if those people also sought healthcare. However, the results were promising with the Apple Watch showing a high true positive rate although some were surprised by the low number of notifications given out. The source is clearly reliable and also highly relevant to this dissertation as these numbers show the strength of the watch and can be used to obtain rough estimates for the economic impact of AF detection software.

[65] This resource shows the population of the USA in 2009. The purpose of the website is to display accurate statistics regarding population and the figures are what one would expect given census data. This piece is important as it helps highlight the prevalence of undiagnosed AF in the USA by using this number alongside the information presented in [47].

[66] This article published by the medicine department of Stanford University summarises the results of the Apple Heart Study, showing a 71 percent positive predictive value (although this number is believed to be higher in reality as paroxysmal AF may go undetected over the seven day period in watch patients were monitored using clinical equipment for signs of AF). Stanford is a reputable University whose research is academically rigorous and reliable and thus likely to be accurate. Once more, this piece is extremely relevant as it analyses the real-world performance of the Apple Watch and enables one to estimate the economic and social impact of such AI technology.

[67] This article, similarly to [66] and [64], summarises the results of the Apple Hearts Study and features an interview with the Apple Heart Study Primary Investigator (an academic working at Stanford University). The information is clearly accurate as it comes from reliable sources, including one of the members who conducted the study. This piece is important as it shows that the Apple Watch is effective at diagnosing AF which is important as it is a feature can be used without the user having to manually take readings. As such, it has massive potential as one only has to wear the watch in order to be screened for the condition. The ability of AI systems to screen for a range of diseases without the participant actively needing to do anything is pivotal as it enables a large group of people to be screened for a wide range of diseases.

[68] This is a website dedicated to calculating the world population. It says that as of January 1st, 2020, the world population was estimated to be 7,763,035,303. The site seems to use a range of reliable sources and the estimate is in line with other estimates. Of course, this is only an estimate and the methodology used is not clear and therefore the number is likely to be off. However, given that other estimates are similar and the range of resources it appears to use in its calculation, I conclude that this estimate is a valid one. This also relevant as it can be used to calculate an estimate for the economic impact of AF detection in the Apple Watch.

[69] This resource has the same publisher as [68]. This source estimates the population of the USA and appears to be using valid and reliable resources to reach its estimate although the actual

method used is unclear. As this estimate aligns with other estimates, I conclude that the estimate is a valid one.

[70] This Statista chart shows the number of smartphones users per year. Statista is a typically trusted website used for its quality statistics (although it is less trustworthy for non-paying members). This estimate of the number of smartphone users aligns with a great deal of other estimates and can be somewhat fact checked through the data published by large mobile phone companies. As Statista only publishes research conducted by its team, whose goal is to give accurate data, the piece is likely to be valid. However, I can not view any of the sources used without purchasing a premium membership which does detract from the reliability of the source. Despite this, I feel the estimate is valid due to the reliability of the author and the fact that other estimates are similar.

[71]. [71] describes information about atrial fibrillation and its stroke risk and states that through treatment, AF sufferers can reduce their stroke risk “by approximately one-half to two- thirds”. [71] is authored by The National Institute of Neurological Disorders and Stroke, a US government agency designed to direct and fund research in the field. Therefore, the purpose of this piece is clearly to inform and present facts regarding the issue, lending validity to the resource. This piece is highly relevant as such a significant reduction in stroke risk through treatment naturally means many lives could be saved and improved through more effective diagnosis of atrial fibrillation.

[72] This Stroke Belt article details the survival rate of stroke sufferers. Notably, the mortality rate within 28 days of a stroke was 28%, a number which increased to 41% after a year. This highlights both how deadly a stroke can be as well as the long term, irreparable consequences that a stroke can cause. The article is summarising the results of a reputable and well cited article produced by researchers at King’s College London and as such, the contents of [72] are both accurate and reliable. Given that [71] highlights the significant reduction in stroke risk amongst AF sufferers who receive treatment, [72] shows that through diagnosis of AF, many lives could be saved. This shows the usefulness of a wearable such as the Apple Watch which allows effective screening of a large population size.

[73] This news article describes the FDA’s approval of IDx-DR, an AI system that detects diabetic retinopathy. Arguably even more exciting is the fact that system has been approved to operate without an expert needing to be present or verify the results. The device determines whether a patient is “negative for more than mild diabetic retinopathy; rescreen in 12 months" or has "more than mild diabetic retinopathy detected: refer to an eye care professional". This technology enables a much larger range of diabetics to be scanned for the condition as no professional is required to analyse each scan initially (which of course means that a far larger number of people are able to be screened for the condition). The systems achieved a roughly 90% accuracy rate. The report is accurate as all of its findings can be verified via the FDA, the IDx-DR website and through the study performed to determine the accuracy of the system. As such the source is valid and relevant as it links into this theme of AI enabling a much larger group of people to be assessed for a disease as it speeds up analysis and requires less professional help. Similar to the Apple Watch, this device ultimately refers people for human healthcare if a condition is detected which is perhaps limitation of the software currently.

[74] This study analyses the effectiveness of IDx-DR (talked about in [73]) to detect mild diabetic retinopathy. Across a modest sample of 900 people, the system achieved a roughly 90% accuracy and these findings are partially responsible for the FDA approval of the system according to the paper. The peer-reviewed entry is clearly accurate as the FDA had to accept the results of the

findings in order to approve the system. However, the confidence in the data is slightly limited by the small sample size but one can safely conclude that 900 people is a large enough sample to show that the system is effective. This ability to accurately diagnose the condition “without the need for a clinician to also interpret the image or results” is a breakthrough in medicinal AI and shows that AI is suitable for replacement of human doctors even if experts are still required to asses more serious cases. This does beg the question as to a possible negative social impact of leading to the redundancy of medical professionals in some fields which leads to unemployment and possibly a stagnation in research. Although, these are arguably outweighed by the ability to screen more people and high levels of accuracy provided by such a system.

[75]. Frick et al. describe the large economic impact of blindness in the USA each year, with an average impact of $5.5 billion spent on medical care and 209000 quality-adjusted life years. A comprehensive survey was performed to obtain data used in this peer-reviewed piece and as such, I have a high level of confidence in the findings of this journal entry. Through effective screening of diabetic retinopathy, the number of cases of blindness could be significantly reduced which would result in large economic benefits (through reduced healthcare expenditure and ability to work effectively) as well as social benefits through less loss in quality-adjusted life years.

[76] This Reuters news article reports on Amazon’s scrapped artificial intelligence system that was used to scan resumes to find the best applicants. The system was trained using Amazon employment data to identify the best candidates. However, Amazon employees are predominantly male which lead the system to conclude that female applicants were inherently far weaker. This piece is published by Reuter, a reputable news organisation, which features interviews from anonymous Amazon employees (who presumably have been verified by the author to actually work for Amazon). Furthermore, Amazon did not deny this to be true but said that the tool “was never used by Amazon recruiters to evaluate candidates.”. Therefore, I conclude that this report is accurate. It is also relevant as it shows one of the massive problems with the black box nature of AI, it is good at training a model to effectively categorise data with relationships but that it may sometimes learn the wrong things and fail to generalise. This is an important concept which could have disastrous consequences in healthcare, if for example medical imagine training data contains a disproportionately high number of white individuals, an AI may falsely assume a correlation between colour of skin and a certain disease. This shows the importance of carefully considering training data, developing methods to understand what an AI uses to make its decisions as well as limiting an AI’s access to irrelevant information e.g. sexuality.

[77] Williams and et al. present their view that black people in America are less healthy than other races and cite racial bias and socio-economic inequity as some of the reasons as to why this is. The piece is somewhat opinionated and anecdotal, highlighting individual cases to support their claims. However, their points are backed up by science and they reference various pieces such as a peer-reviewed JAMA piece which finds that racial bias is partly responsible for the worser health state of black Americans. It also mentions a study across the 171 largest cities in the USA that concludes that white people “living in the worst conditions in urban areas – in terms of poverty rates and single-parent households – are nonetheless residing in circumstances much better than those of the average black person.”. As a result of this, the piece presents a compelling case for medicinal racial inequity in America and can be trusted, although the piece is also not neutral (i.e. it is clearly arguing solely from the side of racial inequity). However, the science agrees with this piece as the picture it paints is consistent with other studies into racial disparity in healthcare. The fear is that artificial intelligence systems may further this as there are fewer black people than white people in America

(potentially leading to software that works better for white people in America), something which has already been seen in facial recognition AI which performs notably worse on people of colour.

[78] This Guardian news article summarises the findings of The Office of National Statistics (a UK government department) that finds that black people are four times more likely to die of COVID-19. It also found that black people are twice as likely to die of COVID-19 than white people when factoring in health, age, and socio-economic factors. This shows that black people are typically worse off when the aforementioned factors are combined. It is still unknown why ethnic minorities die disproportionately more when exposed to the disease but because this gap in death rate is also seen amongst other demographics such as Pakistani , Indian and Chinese households, one can infer that racial bias or inequity may be a large factor in this. These figures are reliable as they are clearly there to inform and performed by a neutral government agency. They show a strong likelihood of unequal health status and / or healthcare in the UK. This once more links back to the idea of AI possibly furthering this gap and also shows the importance of treatment methods that do not enhance the gap in quality of healthcare treatment between those of various races, genders, and sexualities.

[79] This analysis of three studies into racial inequality in healthcare concludes that there is a racial disparity in healthcare although the piece goes on to find that more analysis is needed to determine what extent of this can be attributed to discrimination and what extent of this can be attributed to socio-economic factors (as minorities are typically less well off, which some argue is a result of systemic racism). The piece thoroughly analyses these reputable studies and as such, I believe these findings to be valid. The author shows that adjusting for perceived discrimination does not remove racial inequity, which the author suggests shows a combination of discrimination and socio-economic inequality. This is important as poorer healthcare can lead to a deterioration in socio-economic wellbeing which is precisely why racial bias in AI could have an adverse socio-economic impact on minorities if not addressed.

[80] This online dictionary defines medicine as “The science or practice of the diagnosis, treatment, and prevention of disease” which is clearly an accurate definition. The dictionary makes use of the Oxford dictionary, one of the most trusted dictionaries, making this source reliable and useful to my dissertation.

[81] This YouTube video series discusses what genetic algorithms are and talks about both the theoretical aspect of such algorithms as well as examples of their implementation with accompanying code. Genetic algorithms can cover a wide range of things and are designed to broadly represent evolutionary mechanisms by prioritising agents that perform best by using their genetic code to create new and (hopefully) better agents. This can be a form of unsupervised learning often used alongside a neural network and these algorithms have proved effective in mastering videogames such as “Flappy Bird”. This piece is published by a renowned computer scientist who has written a book discussing the topic and the information is accurate as it aligns with other resources and because the implementation of his algorithms function as expected. This source is relevant as it shows the fact that there are a variety of training mechanisms for artificial intelligence systems, many of which do not involve supervised learning (which is useful given that the medical industry has such a large amount of data that is not conveniently labelled for a computer).

[82] This piece describes the three most common types of training categories for machine learning systems: supervised (use of pre-labelled data), unsupervised (data that has no labels) and reinforcement (navigating a model such as a game board and determining whether the actions it took led to a good result). These principles are common knowledge and can be verified by a great

deal of machine learning resources. This is relevant as it shows that AI systems in healthcare can be trained in many different ways, allowing for diverse systems that can do many things and does not just restrict AI systems to use of labelled data sets.

[83] Singh's article describes AI's capabilities in image recognition including products such as Google Vision and describes how they are capable of facial recognition, detecting objects, detecting text and landmark detection. Singh backs these examples up with various AI systems that can perform such tasks and when this is considered alongside other prominent AI systems, the contents of this article are accurate and valid. AI's competence at image recognition is important as much of medicine uses imaging technology such as MRI scans to detect various conditions, which is an image recognition task. This shows the massive potential of AI in healthcare, something which is reflected in the fact that roughly half of FDA approved AI are imaging AI [Figure 3].

[84] This article details the release of OpenAI's first natural language processing system, a bot that can respond to human input and answer things such as trivia questions or perform basic mathematics. The GPT-3 model has over 175 billion parameters. The contents are accurate as this release was widely reported, the quotes from OpenAI co-founder are accurate and because the product described is available through an application programming interface. The ability of AI to understand and meaningfully respond to human text could prove to be important in seeking to, for example, replace / co-exist with general practitioners. NLP can be used in many parts of health care such as speech bots like Apple's Siri or in analysis of notes taken by doctors to extract meaningful data.

[85] This peer-reviewed nature journal entry talks about many of the challenges that need to be overcome for widespread AI implementation in healthcare to become a reality. This piece particularly focuses on issues of data ownership and lack of investment in the digital architecture needed to create such an environment. The piece stems from a respected journal and uses many reliable resources to make its points. Furthermore, the conclusion it draws is a valid one and it fairly demonstrates that large steps are needed before AI implementation is a universal healthcare reality. This current limit in infrastructure is important as it shows the limited effectiveness of research as it is harder for these AI systems to transition to implementation and also that research is less cost-effective due to lack of data collection protocol.

[86] This book by Thompson discusses various elements of coding and particularly what the lives of coders are like. Thompson dedicates an entire chapter to artificial intelligence and talks about many things such as the success of an AI at playing the game go using reinforcement learning and a less common search tree (the Monte Carlo search). The chapter features an interview with head of AI at Google, Jeff Dean, and details many techniques such as neural networks. The piece is extremely well researched and uses a highly reliable and knowledgeable source (Jeff Dean) and aligns with general knowledge of the subject, making it valid and reliable. The versatility of application for AI as well as its ability to perform with super-human intelligence is impressive and shows promise for application in all fields, including healthcare. The description of neural networks furthered my understanding of the topic and enabled me to understand the strengths and limitations of such systems, something which is transferrable when analysing the pros and cons of AI in healthcare (making concepts such as racial bias in training data leading to flawed systems easier to understand).

[87] This news article by Apple states that their atrial fibrillation detection AI used in the Apple Watch has been approved by the FDA with a "De Novo classification", meaning that the system is available over the counter. The article is biased as it is written by Apple with the intention to paint them in a positive light but the article is relevant to the dissertation only so far as to confirm that the

feature has been proved by the FDA (which is also independently verifiable). This shows that the FDA views the software to be safe and appropriate for use, demonstrating confidence in its accuracy. It also is important as it shows the ability of AI systems to receive governmental certification and essentially shows that the results from the Apple Heart Study can be used to estimate the potential economic impact of the AF detection in the smart watch.

[88] Husten summarises concerns over the “hype” surrounding the atrial fibrillation detection technology in the Apple Watch. The piece lacks neutrality and is largely biased and opinionated with Husten arguing that the Apple Heart Study will lead to lots of anecdotes but little useful data. He argues that the fact that the trial was not a randomized controlled trial is a limitation of the study, which is certainly a valid critique and one of the largest flaws with the trial. However, I feel that Husten’s views are not entirely fair as the large sample size allows for a relatively high degree of confidence in the performance of the Watch. Furthermore, the view that the study led mostly to anecdotes is unfair, in my view, as a large amount of useful data was obtained such as its specificity and sensitivity as well as the percent of cases it identified compared to the number of cases one have may expected it to identify. As a result of this, I feel that most of the critiques are invalid or unfairly negative, but I also feel that some of these viewpoints are valid. The trial was not perfect as it was not a randomised control trial, because it relied on self-reporting, because it relied on self-enrolment (potentially leading to a more digitally literature and hence probably younger sample) and because the study failed to meet its target amount of participants older than 65 (something which is important as AF is far more prevalent in those who are older). From this source, I took away that there were flaws with the design of the Apple Heart Study, even if I perhaps reached a different decision to Husten as I believe that the large sample size of the trail has led to results in which one can be rather confident in. The FDA’s approval of the system strengthens my confidence in the findings of the trial.

[89] This documentary documents the development of an artificial intelligence system named AlphaGo that learned to play the board game Go (a game typically deemed impossible for computers to play well) and beat the world’s best Go player. The system used a reinforcement learning algorithm to master the game as well as a Monte Carlo based tree search and a neural network to prune the search tree to play the game Go. What was so impressive about this was that the system was entirely self-taught and learned only through playing games against itself. This versatility and range of machine learning techniques is impressive and suggest that techniques beyond supervised learning can be applied to the field of medicine. The ability of AlphaGo to beat the Go world champion demonstrates the ability of AI to massively outperform humans even in techniques where its brute forcing skills are far less effective. It essentially shows that AI is applicably to a broad range of things that one may never expect an AI to be able to perform well at, including many tasks in medicine.

[90] This paper by Silver et Al. describes the development of a general reinforcement algorithm that masters a range of board games such as Chess, Go and Shogi. The ability to apply a general algorithm to such a diverse range of complex games shows the promise of such techniques particularly as it massively outperformed the best chess player (including human and non-human players) with only a few hours of training. Through clever search techniques such as Monte Carlo search, it was able to beat Stockfish (the best chess bot) by analysing only 60,000 board states compared to Stockfish’s 60 million per turn. Although reinforcement techniques have yet to be applied to the field of medicine, the promise of such algorithms is clear given their ability to perform so well in games that were previously believed to be impossible for computers to play to such a standard.

[91] This documentary looks into the ethics surrounding artificial intelligence systems. The piece discusses a wide range of ethical considerations regarding AI use and particularly highlights the lack of legislation regarding ethical AI and the incomplete nature of governmental guidance on ethical AI. It talks about the threat of AI in a range of things. The documentary is published by a German public broadcast service, DW, on their YouTube channel dedicated to documentaries. The piece is therefore likely to be well researched and neutral. A large amount of the documentary is based on interviews with certain experts which could lead to a biased view of the overall ethicality of AI as the people being interviewed are focusing solely on the ethical disadvantages of AI, although these downsides are clearly still valid. The piece is focused on broader AI use as opposed to just medicinal AI which limits the relevance of the source. However, this concept of weak legislation, weak ethical guidance and potential to cause serious harm are somewhat relevant in showing that AI as a whole has many hurdles to overcome, particularly if it is to be implemented in a life-saving industry such as medicine.

[92] This 2015 newspaper article talks about tech companies investing into Artificial Intelligence, particularly in its advanced natural language processing capabilities. This news article from a South Carolina paper is broadly accurate as the information algins with publicly available information such as Google acquiring companies such as DeepMind and with my own knowledge. However, the article is limited by its simplistic description of the field which is not highly informative. It is definitely useful to the dissertation though as it highlights the high level of investment into AI in 2014, showing that as recently as 2014, a large amount of AI was still being researched as opposed to implemented. It also has use as it highlights the advanced capabilities of NLP, an area of AI that many are currently using to research healthcare bots.

[93] This comprehensive review of AI systems that have been researched with the potential of being used to treat diabetes concludes that AI has the potential to revolutionise the field as it offers greater accuracy and efficiency than current protocol. The piece is a thorough assessment of all relevant research into diabetic diagnosing and treating AI, with over 450 studies looked at including the most notable ones being analysed to some depths such as IDx-DR (described in [74]). This piece uses a range of reliable and open resources and is a comprehensive and valid analysis of the current state of AI in the field of diabetes and as such is highly valid. The piece shows the wealth of research occurring and the comparatively far fewer number of FDA approved uses as well as the ability of AI to generally outperform human doctors in various fields.

[94] This YouTube video features members of the Montefiore Medical Center of the Albert Einstein College of Medicine discussing their use of AI in their practice. They highlight the extremely large amount of data they have to deal with which they say is advantageous as it allows them to train highly effective AI models. The resource is flawed as the video was produced in partnership with Intel, a computer company, whose healthcare team is working with the hospital members to develop the systems. As such, this video will almost certainly paint their systems in an overly positive light detracting massively from the validity of the source. There is still some validity as these professionals clearly are benefiting from AI in the field and do know what they are talking about. This idea of larger data sets leading to more accurate models is important in the context of medicine as hospitals typically store large amounts of data on the medical history of their patients which would facilitate more effective models.

[95] This piece by Google discusses the importance of data sets when training a machine learning model, highlighting that a data sets of high quantity, high accuracy, and minimal skew (i.e. representative data) are traits that are needed to produce the most accurate models. Google implies that more complex systems are only feasible with relatively large data sets. The piece is clearly

intended at developers as it is part of a machine learning course produced by Google, meaning that the information presented is accurate and reliable. As it aligns with my own knowledge and other resources, I conclude that the information presented here is valid. This idea of the need for large data links back to a recurring theme discussed in [94] and [98] and the concept of a “skew” in training data being bad links to theme of potential bias (such as racial bias) in AI systems as discussed in [76].

[96] This economist video assesses the potential future of AI in healthcare and features an interview with a pioneer in the field. Given the reputation of the economist, the qualification of Dr Eric Topol (who published [Figure 2] and [Figure 3]) and the fact that the information presented aligns with other resources, I conclude that this source is accurate and valid. This piece crucially highlights a key advantage of AI: it does not get tired and always preforms consistently. It is able to perform analysis tasks in milliseconds which might have taken a human 15 minutes, which is important in things such as imaging as faster analysis enables more scans to be performed. For something like breast cancer, this would allow a greater range of people to be scanned more frequently, something which would save lives.

[97] This Bloomberg documentary assesses the rise of artificial intelligence with a particular focus on its birthplace of Canada. The purpose of the piece appears to be a mix between to inform and to entertain (as shown with its rather informal and comedic tone). The piece is also likely to be somewhat biased as it features an interview with Justin Trudeau (the current prime minister of Canada) who is talking about Canadian contribution to AI. However, he does raise the importance of ethics in AI, something which is arguably most vital in the medicine industry. Whilst the contents of the source are accurate, they are mostly irrelevant and therefore I have chosen not to use this piece in my dissertation. However, the more general points about AI have been useful in understanding other resources.

[98] This 2019 panel amongst medicinal AI industry leaders discusses the current state of AI in healthcare and what issue need to be overcome (and how) in order for this implementation to become widespread. Various use cases such as image analysis, extracting meaningful data from notes taken by doctors and general diagnosis are discussed. The piece focuses on a large theme of this dissertation (which is discussed in sources such as [35]) and one of the primary problems with implementation of AI into healthcare: its black-box nature. One interesting point raised by Doctor Katherine Andriole is that the black-box nature is still a large problem that must be overcome for widespread implementation to be possible but that techniques such as a saliency map have been successfully developed to somewhat mitigate this. This technology can, for example, show to doctors which features of a patient’s image are being used to calculate the decision. These maps have been used to show that, in some cases, AI analyses the exact same features a human doctor would which she says allows for a greater level of trust in the systems. However, all the panellist including her conclude that this alone does not suffice as they feel that if a doctor can not trust the system, then the patient can not either. This is however a contentious issue and some feel that the black-box trade-off may be worthwhile given better accuracy.

[99] This piece discusses the creation of a rule-based healthcare system that encodes knowledge of various heart doctors to produce clinical decisions based on factors such as blood pressure. The report concluded that the system led to improved quality of life. The piece shows the potential for rule-based AI systems to be effective. The piece is published by reputable authors and the methodology is sound lending validity to its findings. This piece also highlights how dependent such AI systems are on human knowledge and as such typically only have potential in the automation of simple tasks. Whilst the piece is valid, I did not directly use this source. I used it to enhance my

background knowledge, but the source was ultimately not included due to the fact that most rule-based systems are very limited in their ability as they depend on human knowledge. This one was more limited as the rules were not widely accepted rules but merely ones drawn up by a small set of doctors, making it a poor case to use to assess the socio-economic impact of such AI systems.

[100] This live streamed Harvard Medical School lecture discusses the future of AI in medicine by looking at current research and use as well as issues to widespread implementation. The piece states that models are" as good as the training data" and highlights the need to address biases in data to ensure truly intelligent models. These are recurring themes which are discussed in a number of resources such as [95] and [76] and are therefore clearly valid issues that are prevalent with medicinal AI. The purpose is clearly to inform and as the lecture is being carried out in one of the most prestigious academic institutions in the world, the contents are reliable. Therefore, I feel the source is of high validity.

[101] This Harvard Medical School Seminar discusses AI in healthcare. The focus is on AI systems that are being (or have been) used in healthcare or are being researched. In short, it analyses functioning healthcare AI. Whilst [100] is much more focused on looking at a general overview of the future of medicinal AI, [101] is a compilation of isolated instances of healthcare AI. The piece highlights the ability of AI to perform at similar and better levels than humans, which is a recurring theme of research and implementation of AI. One talks about pattern recognition and another looks at an AI that analysed blood types to determine appropriate treatments. The speakers and the academic institution hosting the speakers are reputable and their anecdotes are verifiable lending reliability to the source. I have not used this piece in my dissertation as the focus is on individual and often non FDA approved systems which are hard to assess when attempting to a look at an overall impact of AI (anecdotes are not as strong as data). Many of these AI systems are also atypical and as such are not broadly representative of most systems in the field. Therefore, the source is valid but has not been directly used.

[102] This talk by Doctor Matthew Lungren talks about AI use in radiology as a whole and discusses both advantages and disadvantages of the field as well as the fear of some that radiologists will become obsolete in the next few decades. Lungren talks about the fact that AI is able to greatly speed up time for scans to be analysed as it is far faster than it is possible for a human to be. The piece also explains what a saliency map does, a concept also discussed in [98], and how it can reduce the issue of the black-box nature of AI. The talk is given by an expert at Stanford, the content of the talk is clearly accurate, and his purpose is to be informative. Therefore, the source is valid, particularly as it is so relevant to the breast cancer AI system that this dissertation evaluates in some depth.

[103] This talk by Ben Glocker details the use of deep learning techniques in healthcare and particularly focuses on neural networks using both supervised and unsupervised learning techniques. The piece shows just how effective AI are at complex image analysis, something which Glocker demonstrates can and is being used in medical imaging analysis. Glocker is a highly qualified professional, giving lectures about medical image computing at Imperial College London and is a senior member in a research group that focuses on applying machine learning techniques to biomedical image computing and medical computer vision. Given his high level of expertise and given that many of the cases he speaks of can be verified, it is clear that this source is highly valid, particularly showing the potential for AI in medical image analysis.

[104] This keynote session talks about how companies such as Google are attempting to utilise machine learning advances in the field of healthcare. The talk uses a range of medical examples to

show the promise of using AI in healthcare, with the theme of artificial intelligence being competent at image recognition running throughout the talk. The talk may be slightly biased as the speakers are associated with Google and their parent company and as such may wish to highlight Google's machine learning library: TensorFlow. However, this is a negligible concern as the two are qualified professionals who use a range of valid examples to make their points, making the source valid. As the talk focuses so much on possible future application, it is slightly irrelevant. Much more relevant is the ability of AI to perform so well at image recognition, a core theme of the dissertation, and the fact that these industry leaders are confident that AI can, as it is at the moment, be successfully applied to medical problems and perform well.

[105] This podcast on AI in healthcare interviews Microsoft Australia's Chief Medical Officer (Doctor Nic Woods) who discusses the general state of AI in healthcare. Doctor Woods talks in detail about how much of the research into medicinal AI is centred around the analysis of medical scans to form a diagnosis. Given the qualification of Doctor Woods and the educational focus of the podcast, I would conclude that this resource is valid even if it may be limited by the fact that it is coming from the perspective of a Microsoft employee who may give a more positive view of the state of AI in healthcare when compared to a radiologist currently working in a hospital, for example.

[106] This piece provides an overview as to what artificial intelligence actually is and provides an overview of its history, general subsets (such as machine learning and rule-based systems) as well as common applications such as natural language processing and image recognition. The piece appears to be designed to be informative even if it does not stem from a highly reputable source. However, the contents of this tutorial are extremely accurate and aligns with general knowledge around the subject. The history of AI segment is very well researched which lends massive validity to this source. This general overview of different types of AI as well as their strengths is directly transferable to medicine (for example this recurring theme of AI being strong at image recognition is something that transfers to fields of medicine such as mammogram analysis).

[107] This in-depth introduction into machine learning (which is the first piece in a wider course) provides a detailed understanding of the processes by which machines learn such as supervised and unsupervised. It is much more in depth than a resource like [106] and is also very relevant as it highlights potential uses of unsupervised learning methods in medical AI (such as the clustering of data, something which may be useful to categorise groups of patients into certain categories). The content of this article aligns with widely accepted AI knowledge and matches a range of other advanced insights into the topic such as [7]. Therefore, the unreliable nature of [107]'s medium is of little importance, particular as the author is a specialist in data science and machine learning. As a result of the accuracy and balanced nature of the resource, it is clear valid.

[Figure 1] This picture on stackexchange is the result of running an open source script produced by Mark K Cowan. It is a picture of a typical neural network architecture. This neural network architecture matches other common drawings of the structure and hterfore is clearly accurute.

[Figure 2] This image is taken from a tweet by Eric Topol, a verified Twitter user and a renowned American cardiologist. This picture is of all FDA approved AI systems as of January 2019. Due to Topol's expertise in the field, this photo is likely to be accurate and well-researched, particularly as it uses a highly reputable Nature article for the basis of this source.

[Figure 3] This is an image in the same tweet as [Figure 2] but this time showing FDA approved AI applications as of April 2019. Unlike, [Figure 2] Topol seems to have compiled the list himself perhaps leading to mistakes. All of the examples listed are accurate (something I have verified myself), although some of the comments on the Twitter post suggest that Topol may have missed a few approvals with one user pointing out three possible cases that Topol may have not included (although some may argue that the examples mentioned do not strictly qualify for Topol's list). Despite this, Topol has clearly done his research and included all, or the vast majority, of FDA approved healthcare systems. Therefore, this source and [Figure 2] are highly useful to the dissertation as they show the small-scale implementation of AI as well as its rapid growth over a few months.

[Figure 4] Figure 4 shows a java application that was coded by the author of the dissertation (Philip Mortimer). This tool helps visualise the performance of a neural network at recognising handwritten digits using the MNIST database. The neural network was also trained and coded by Mortimer. This tool can be accessed via the link provided. As, I produced the piece, I know it is accurate and valid. It helps prove a theme of this essay: that AI is highly effective at image analysis.

[Figure 5] An estimate for the potential economic impact of the Apple Watch is provided in [Figure 5] by combining data from the Apple Heart Study, the prevalence of AF and the economic impact of AF. The mathematics is explained in the source. However, the figures only serve as rough estimates as there are a range of factors that the piece simply can not account for due to lack of information. The piece is limited by its simplistic methodology but the ability of the Apple Watch to save around $7 (when only looking at healthcare expenditure) per US citizen who has irregular pulse notifications enabled and does not suffer from AF, is an important one. This figure may well be greater as it does not account for potential income lost and a better quality of life. The figures do show the potential of the watch to have a large economic impact even if one should only interpret it as a rough estimate that is probably far off the truth. That being said, I do have confidence that I performed the mathematics correctly and that the methodology is valid (even if it is simplistic) and that its finding of a significant economic impact of the Apple Watch's AF system is an important and valid one.

[Figure 6] This picture shows an example of the Apple Heart App used alongside the Apple Watch. This picture is clearly accurate as it comes directly from Apple and can be verified through a simple Google search. As such the piece is valid and helps provide a sense as to what the technology looks like.

[Figure 7] This picture shows an example of an irregular pulse notification from an Apple Watch and also comes from Apple, clearly making it accurate. This piece is relevant as it helps provide a greater understanding of the technology being discussed by providing a clearer image of how it actually functions.