The Utilization of Artificial Intelligence in Healthcare and Its Accuracy in Medical Decision-Making

The Utilization of Artificial Intelligence in Healthcare and Its Accuracy in Medical Decision-Making

Julie T. Joseph

A Thesis in the Field of Bioinformatics

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2022

Abstract

Artificial intelligence (AI) is a branch of computer science that involves the use of machines to simulate human intelligence, such as learning and problem-solving. AI encompasses machine learning and natural language processing. Machine learning is a branch of AI that involves the use of machines to learn from new data to improve its predictive accuracy over time. Natural language processing is a branch of AI that involves the use of machines to read, understand and derive meaning from human language, such as speech or free text. The role of AI is evolving and expanding in healthcare from its use in disease screening tests, disease diagnosis, predicting survival prognosis and predicting disease complication.

Healthcare has its own set of unique functional and ethical challenges. Unlike other industries that utilize AI, the healthcare industry’s utilization of AI is limited due to very large datasets that are siloed and not standardized. These factors prevent interoperability and consequentially compromise patient safety and the quality of medical care. Although, there are some potential risks and challenges associated with the utilization of AI in the healthcare setting, the benefits appear to outweigh the risks.

Dedication

This thesis is dedicated to my family who have supported me throughout my education. I thank them for their endless love, support and encouragement of lifelong learning.

Acknowledgments

I would like to acknowledge and give special thanks to my thesis director, Theodore Hatch Whitfield, who guided and advised me throughout the process of writing my thesis.

Table of Contents

Chapter I.

Introduction

Background

Artificial intelligence (AI) is a branch of computer science that involves the use of machines to simulate human intelligence, such as learning and problem-solving. AI encompasses machine learning and natural language processing. Machine learning is a branch of AI that involves the use of machines to learn from new data to improve its predictive accuracy over time (IBM). It finds and applies patterns in data. Machine learning was coined by Arthur Samuel in 1959 as "the field of study that gives computers the ability to learn without explicitly being programmed" (Chowdhury, Apon, & Dey, 2017). Natural language processing (NLP) is a branch of AI that involves the use of machines to read, understand and derive meaning from human language, such as speech or free text. Written text, radiology images, videos, audio and faxes are some examples of unstructured data. Unstructured data is making it difficult to be fed into algorithms. NLP processes and analyzes unstructured data to gain meaningful insights (IBM).

John McCarthy, a professor emeritus of computer science at Stanford University and who is widely recognized as the father of artificial intelligence, first coined the term "artificial intelligence" in 1956 at the Dartmouth Summer Research Project in Artificial Intelligence conference. Other "founding fathers" of artificial intelligence include Alan Turing, Marvin Minsky and Allen Newell. Alan Turing was an English mathematician and computer scientist who developed the Turing machine which provided the algorithm that led to general-purpose computers. Marvin Minsky was an American cognitive and computer scientist who co-founded the Massachusetts Institute of Technology AI Laboratory in 1959 with John McCarthy. Allen Newell was a computer science researcher who developed two of the earliest AI programs called the General Problem Solver and the Logic Theory Machine (Solutions).

The role of AI is evolving and expanding in healthcare from its use in disease screening tests, disease diagnosis, predicting survival prognosis and predicting disease complication. Does the use of AI in hospitals result in improved precision and accuracy in medical decision-making and is it worth the investment to utilize AI in the healthcare setting? What are some of the obstacles and risks associated with AI implementation in the healthcare setting?

According to Davenport and Kalakota, "It seems increasingly clear that AI systems will not replace human clinicians on a large scale, but rather will augment their efforts to care for patients" (Davenport & Kalakota, 2019). Doctors will find changes to, rather than replacement of, their job.

AI will be used to support clinical decision-making. Clinical decision-making involves an emotional component and socioeconomic component. The emotional dynamics of human interaction in healthcare decision-making is a limiting factor that prevents AI from replacing clinicians. The discussion of a medical diagnosis or a treatment plan is based on a relationship between a physician and a patient that is built on trust (Review). The physician must be able to empathize with the patient and respond with compassion. In addition, diagnosis and treatment decisions must take into consideration social factors such as living conditions, family support system and access to care. Patients living alone without social support or cognitively impaired patients who are unable to comprehend medical information or lack the capacity to make healthcare decisions need to be taken into consideration when making clinical decisions. AI is unable to process the emotional and socioeconomic factors that are an essential in the clinical decision-making process. Thus, AI will not likely replace medical doctors by and large.

Patients generate a ton of data and this data is used by hospitals and physicians to improve the accuracy, quality and speed in providing medical care. Machine learning can build better patient profiles and predictive models to effectively diagnose and treat patients. NLP can analyze unstructured medical notes such as free text data, physician order data or dictation notes and provide insights and summarize information. NLP uses two analytic techniques called syntactical analysis and semantic analysis. Syntactical analysis uses the rules of grammar to examine the correctness of the sentence. Semantic analysis extracts logical meaning to the sentence (IBM). According to Bresnick, "In the healthcare industry, natural language processing has many potential applications. NLP can enhance the completeness and accuracy of electronic health records by translating free text into standardized data" (Bresnick, 2016). Several research studies conducted on the accuracy of machine learning and NLP in medical decision-making have reinforced the higher accuracy and precision of AI in clinical decision-making.

Kelly and colleagues state "analysis of the immense volume of data collected from electronic health records offer promise in extracting clinically relevant information and making diagnostic evaluations as well as in providing real-time risk scores for transfer to intensive care, predicting in-hospital mortality, readmission risk, prolonged length of stay and discharge diagnoses, predicting future deterioration, including acute kidney injury, improving decision-making strategies, including weaning of mechanical ventilation and management of sepsis" (Kelly, Karthikesalingam, Suleyman, Corrado, & King, 2019). In addition, AI in healthcare can help save time and resources by streamlining tasks.

Functional and Ethical Challenges in AI in Healthcare

AI faces many functional and ethical challenges, primarily revolving around interoperability. Interoperability is the ability of health information systems in various healthcare entities (hospital, physician/provider, health insurance) to seamlessly exchange patient data across organizations to advance the effective delivery of healthcare (HIMSS, 2021). Interoperability is a key factor for AI to work effectively. The healthcare system involves a number of electronic health record (EHR) systems that contain patient data. A patient can have their inpatient, outpatient, laboratory results and radiological exam data in four different EHR systems that do not communicate with each other. In order for AI to work efficiently, all EHR systems of a patient's data need to communicate and exchange information flawlessly. Unfortunately, this is not the case in the current United States (U.S.) healthcare system.

Historically there was have been a number of impediments to achieving interoperability across the U.S. healthcare system. The lack of universal adoption of a standardized EHR system, the complex privacy and security challenges, the impact on health providers' workflow and weak or misaligned incentives are some historical impediments to achieving interoperability in the U.S. healthcare system (Reisman, 2017).

Fragmentation of the EHR system along with competing standardized formats among EHRs have resulted in incompatible information exchange and lack of interoperability. Privacy and security laws vary from state to state making cross-state health information exchange and interoperability difficult (Mello, Adler-Milstein, Ding, & Savage, 2018). The workflow processes providers follow to care for patients are not standardized and when these workflows are redesigned to allow for interoperability, it causes a significant negative impact on a provider's day-to-day workflow. In addition, historically there were no incentives provided by the government or payers to promote interoperability or widespread information exchange with external organizations outside a provider's health system (Campbell, 2015).

Chapter II.

Research Methods and Limitations

A review of the current English literature in PubMed revealed a number of studies on the use and accuracy of AI technology in healthcare decision-making, specifically pertaining to disease screening tests, diagnosing disease, predicting survival prognosis and predicting disease complication. Keywords applied include artificial intelligence, machine learning, natural language processing, screening test, disease diagnosis, disease complication, survival prognosis prediction and accuracy rate. These peer-reviewed studies were summarized and the accuracy of AI application in medical decision-making was evaluated. The accuracy of AI in medical decision-making was further explored in both physical health and mental health disease. For example, the AI accuracy of a radiological exam in cancer screening and the identification of symptoms of severe mental illness are some of the studies that have been reviewed.

Evaluation of the precision and accuracy will be analyzed through the sensitivity, specificity, area under the curve-receiver operating characteristic curve (AUC-ROC) and the accuracy rates published in these studies. True positive is the number of cases correctly identified as a positive for disease. False positive is the number of cases incorrectly identified as positive for disease. True negative is the number of cases correctly identified as negative for disease. False negative is the number of cases incorrectly identified as negative for disease (Baratloo, Hosseini, Negida, & El Ashal, 2015).

Sensitivity, also known as the true-positive rate, is the proportion of those who test positive out of the those who actually have the disease. A highly sensitive test would imply there are few false negative results and fewer cases of diseases missed. Specificity, also known as the true-negative rate, is the proportion of those who test negative out of the those who do not actually have the disease. A highly specific test would imply there are few false positive results (Safari, Baratloo, Elfil, & Negida, 2015).

Accuracy is the closeness of a measurement to the true value. Accuracy is the proportion of true positive results (true positive and true negative) in the total selected population. Accuracy is equal to (sensitivity) (prevalence) + (specificity) (1 –prevalence). The area under the curve-receiver operating characteristic curve (AUC-ROC) value is a very important classification model performance measurement in machine learning (Narkhede, 2018). The higher the AUC-ROC is the better the model is.

The focus will be on reports assessing the use of AI in the inpatient (hospital) setting. Outpatient (ambulatory care, clinic, urgent care) settings will be excluded. In addition, the functional and ethical challenges, as well as the risks and benefits associated with AI in healthcare will be reviewed and discussed. For the purpose of this thesis, two specific branches of AI, that is machine learning and natural language processing, will be assessed. An internal limitation is that supplemental branches of AI such as deep learning will not be reviewed. Deep learning, also known as deep neural network, is a branch of AI which uses layered algorithms to analyze data. Deep learning mimics the human brain with algorithms of three or more layers. Bresnick states, “In deep learning models, data is filtered through a cascade of multiple layers with each successive layer using the output from the previous one to inform its results” (Bresnick, 2018b). Deep learning differs from machine learning in that machine learning requires pre-processing of the hierarchy of features determined by a human expert while in deep learning there is no need of pre-processing of the data. An external limitation is that literature on healthcare AI published in languages other than English will not be included.

Chapter III.

Results

AI uses complex algorithms to emulate human cognition to analyze and interpret medical data. “There are a number of research studies suggesting that AI can perform as well as or better than humans at key healthcare tasks, such as diagnosing disease” (Davenport & Kalakota, 2019). AI can be useful not only in diagnosing disease, but in screening for disease, in predicting survival prognosis and in predicting disease complication.

Screening tests are conducted to detect disease in individuals who are asymptomatic. AI is increasingly being evaluated in its role in identifying disease on screening tests as well as in its accuracy in diagnosing disease. Additionally, AI is being assessed in its accuracy in predicting survival prognosis and the development of disease complication.

AI Accuracy in Disease Screening Tests

The U.S. Preventative Services Task Force (USPSTF) is an independent group of preventative medicine experts that makes evidence-based recommendations on clinical preventative services such as screening tests. Screening tests are done to identify medical disease in asymptomatic individuals providing the potential opportunity for early intervention of an identified medical condition.

Breast cancer

According to the American Cancer Society, “Breast cancer is the most common cancer in American women and the second leading cause of cancer death in women” (Society, 2021a). Some risk factors include age over 50 years, inherited mutations, having menses before the age of 12 or starting menopause after 55 years (C. f. D. C. a. Prevention).

A mammogram is an x-ray of the breast and it is the recommended test for breast cancer screening (C. f. D. C. a. Prevention). The USPSTF recommends screening mammograms every other year in women between the ages of 50 and 74 years. Individuals outside this age range and with certain risk factors for breast cancer are recommended to discuss the need for a screening mammogram with their primary care physician (Force). Mammograms are interpreted by a radiologist, a medical doctor that interprets medical imaging tests such as x-rays, computed tomography (CT) scans and magnetic resonance imaging (MRI). The experience of a radiologist in reading a mammogram as well as the quality of the image are crucial to accurately read a mammogram.

A recent study published in Nature demonstrated that an AI system outperformed radiologists in detecting breast cancer using screening mammograms at two screening centers, one in the U.S. and the other in England. The AI system produced 5.7% fewer false-positives and 9.4% fewer false-negatives than radiologists at the U.S. center, while the AI system produced 1.2% fewer false-positives and 2.7% fewer false-negatives than radiologists at the UK center. At the U.S. center, a separate analysis of 500 randomly selected screening mammograms was done and it revealed the AI system performed better than six radiologists. The AUC-ROC for the AI system was greater than the AUC-ROC for the average radiologist by an absolute margin of 11.5% (McKinney et al., 2020).

However, Schaffter and colleagues showed that the utilization of an AI algorithm with a radiologist evaluation to be more accurate. They conducted a diagnostic mammography accuracy study that involved 144,231 screening mammograms between September 2016 and November 2017. They found that combining an algorithm with a radiologist assessment had a higher sensitivity (92.0%) than utilizing the AI algorithm alone (66.2%) (Schaffter et al., 2020).

If a screening mammogram is indeterminate, an invasive procedure such as a breast tissue biopsy may be needed. The increased accuracy rate of AI in the screening of breast cancer could allow for detection and treatment at earlier stages of cancer. In addition, unnecessary and invasive procedures such as a breast tissue biopsy could be avoided.

Lung cancer

The American Cancer Society reports lung cancer as the second most common cancer among men and women in the U.S. (Society, 2021c). The USPSTF recommends annual lung cancer screening with a low-dose computed tomography (LDCT) in adults between the ages of 50 and 80 years who have a 20 pack-year smoking history and currently smoke, or have quit smoking in the last 15 years (Force, 2021). A LDCT is typically read by a radiologist who provides their subjective interpretation of the image.

Lung cancer typically present with a persistent or worsening cough, chest pain associated with deep breathing, hoarseness of the voice and shortness of breath. Cigarette smoking and exposure to second hand smoke, radon and asbestos are risk factors to the development of lung cancer (Society).

Joy Mathew and colleagues identified 39 studies conducted between 2010 and 2020 on the role of AI in lung cancer screening and found a higher sensitivity, specificity, and accuracy of lung cancer screening and classification of nodules through the utilization of AI in radiological imaging (Joy Mathew, David, & Joy Mathew, 2020). In one of the studies, the pulmonary nodule classification accuracy rate was 91.2 %, supporting AI as a potentially valuable tool in the detection of lung cancer.

If a screening LDCT is indeterminate, an invasive procedure such as a lung tissue biopsy may be needed. The increased accuracy rate of AI in the screening of lung cancer could allow for detection and treatment at earlier stages of cancer. In addition, unnecessary and invasive procedures such as a lung tissue biopsy could be avoided.

Helicobacter pylori

Helicobacter pylori (*H.plyori*) is a bacterial infection of the stomach that is a known risk factor for gastric cancer if left untreated (T. C. Clinic). AI is being utilized in gastroenterology, a branch of medicine focused on the digestive system, in helping screen for *H.plyori* infection using images from an endoscopy. An endoscopy is a flexible tube that is used by a gastroenterologist, a medical doctor who specializes in the human digestive system, to examine the digestive tract.

Bang and colleagues identified eight studies that had a total of 1719 patients (385 patients with *H.pylori* infection and 1334 controls) (Bang, Lee, & Baik, 2020). They used AI algorithms to screen for *H.pylori* infection based on endoscopic images. The area under the curve of AI for the prediction of *H.pylori* infection was 0.92 and the accuracy rate was 82%. The pooled sensitivity and specificity of AI prediction of *H.pylori* infection based on endoscopic image recognition was 87% and 86% respectively.

Definitive diagnosis of *H.pylori* is by an endoscopic biopsy, which is an invasive procedure (T. C. Clinic). The increased accuracy rate of AI in the screening of *H.pylori* based on endoscopic image recognition could avoid unnecessary and invasive tests such as an endoscopic biopsy. In addition, it could allow for earlier detection and treatment of *H.pylori* infection.

Mental Health

Although AI is more prevalent in medicine for its application in physical health, its application in mental health is increasing. Mental health screening is typically done by a primary care physician during the annual physical exam. Patients are questioned about symptoms of depression, anxiety and thoughts of suicide. The Personal Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-2 (GAD-2) are screening tests for depression and generalized anxiety, respectively (Kroenke et al., 2016). According to the American Foundation for Suicide Prevention, suicide is the tenth leading cause of death among Americans and on average there are 130 suicides per day (A. F. f. S. Prevention).

Graham and colleagues reviewed 28 original research studies on AI and mental health. NLP techniques were used to read the raw data from EHR, mood rating scales, brain imaging data, novel monitoring systems (smartphone, video) and social media platforms such as Twitter. NLP analyzed data in the form of text such as clinical notes and conversation such as counseling sessions. AI revealed high accuracy rates in identifying symptoms of severe mental illness (90%) and in predicting suicidal ideation (92%) (Graham et al., 2019).

The high accuracy rate of AI in screening for mental health symptoms indicative of psychosis or suicide could potentially prevent psychotic episodes as well as suicide.

Diabetic Retinopathy

According to the American Diabetes Association, 10.5% of the U.S. population has diabetes (Association, 2021). Diabetic retinopathy is a complication of the diabetes that affects the eye. Diabetic retinopathy is a weakening of the vessels of the eye that supply a part of the eye called the retina. Screening for diabetic retinopathy is done through a dilated eye exam. Diabetic retinopathy is an eye condition that can lead to vision loss and potentially blindness (Institute).

Screening and diagnosis of diabetic retinopathy is via a dilated eye exam which examines the fundus of the eye. Diabetic retinopathy is a weakening of the vessels of the eye that supply a part of the eye called the retina. The retina is inside the fundus of the eye. A dilated eye exam is done by an ophthalmologist, a doctor who is an expert on medical conditions of the eye (T. C. Clinic).

Professor He and colleagues utilized AI algorithms on dilated eye exam photos of the fundus to detect diabetic retinopathy in 889 patients who visited PengPu Town Community Hospital in Jing'an district in China between May 30, 2018 and July 18, 2018. All patients had a diagnosis of diabetes. The AI screening algorithm showed a high sensitivity (90.8%), specificity (98.5%) and AUC (0.95) in predicting diabetic retinopathy in diabetic patients (He et al., 2020).

Mild diabetic retinopathy does not require medication or an invasive procedure. However, if the dilated eye exam shows advanced diabetic retinopathy, then invasive treatments are needed such as an injection of medication through a needle, a laser treatment called photocoagulation, or a surgical procedure called vitrectomy (Institute).

The increased accuracy rate of AI in the screening of diabetic retinopathy could avoid unnecessary and invasive treatments such as the injection of medication, laser photocoagulation or surgery such as vitrectomy. In addition, it could allow for earlier detection and treatment of diabetic retinopathy before it progressively worsens.

Abdominal Aortic Aneurysm

An abdominal aortic aneurysm is an enlargement of the aorta, the main blood vessel that supplies blood to the abdomen, pelvis and lower extremities. An abdominal aortic aneurysm typically grows slowly and without symptoms. It can be life-threatening if the aneurysm bursts and the only curative treatment is surgical repair (T. M. Clinic). According to the CDC, in 2019, 59% of deaths in men were due to aortic aneurysms (Prevention, 2021). The main complication of untreated abdominal aortic aneurysm is a ruptured aneurysm leading to death. Abdominal aortic aneurysms are more common in men, smokers and people aged 65 years and older (T. M. Clinic).

Screening and definitive diagnosis of abdominal aortic aneurysm is by abdominal ultrasound. The USPSTF recommends a one-time screening ultrasound in men who have smoked and are between the ages of 65 and 75 years (Force). Treatment could either be observation or surgery. The need for surgical intervention is dependent on a number of factors including the size of the aneurysm (T. C. Clinic).

Canchi and colleagues assessed 312 patients diagnosed with abdominal aortic aneurysms to evaluate the accuracy of AI in predicting the likelihood of aneurysm rupture based on abdominal radiological images (Canchi, Ng, Narayanan, & Finol, 2018). Four machine learning classifiers used include decision trees, naïve Bayes, logistic regression and support vector machines. AI accuracy rate was noted at 95.2% with an AUC of 0.974.

The increased accuracy rate of AI in the screening of abdominal aortic aneurysm could avoid unnecessary and invasive procedures such as surgical repair. In addition, the increased accuracy of the AI screening algorithm could also help evaluate the indication and need for surgery as well as facilitate the development of a personalized therapeutic approach in patients with abdominal aortic aneurysms.

Skin Cancer

According to the Skin Cancer Foundation, skin cancer will be diagnosed in one in five individuals by the age of 70, and more than 9,500 people are diagnosed with skin cancer daily (S. C. Foundation). Skin cancer screening involves a total body physical examination of the skin and a biopsy of the skin tissue in suspected cancerous-looking skin lesions. This is typically done by a primary care physician or a dermatologist, a specialist of the skin. If cancer is detected, then surgical removal of the cancer along chemotherapy and radiation may be needed (T. M. Clinic).

Skin cancer is an abnormal growth of skin cells and it commonly develops secondary to exposure to the sun. As a result, skin cancer occurs in areas commonly exposed to the sun such as the face, hands and scalp. The three main types of skin cancer are melanoma, basal cell carcinoma and squamous cell carcinoma.

Giavina-Bianchi and colleagues developed AI algorithms for melanoma screening in the primary care setting. They conducted clinical (taken with a smartphone) and dermoscopic (taken with a dermatoscopy device) image classification on over 30,000 images. The AI model proved a higher accuracy rate of 92%, sensitivity of 96%, specificity of 98% than the dermatologists' diagnosis (Giavina-Bianchi et al., 2021).

The increased accuracy rate of AI in the screening of skin cancer could allow for detection and treatment at earlier stages of skin cancer. In addition, unnecessary and invasive procedures such as a skin tissue biopsy could be avoided.

Child Abuse

According to the National Children's Alliance, almost 700,000 children are abused annually and an estimated 1,770 children die from abuse in the United States (Alliance). Child abuse can include sexual abuse, physical abuse, neglect and witness to violence. Diagnosis is made by evaluation of the child's medical history, physical examination, laboratory tests and imaging studies such as x-rays. Abused children have a higher mortality than non-abused children (Shahi et al., 2020) .

Shahi and colleagues identified 737 Level 1 trauma pediatric patients who were abused. The NLP model prove a higher accuracy rate of 93.4%, sensitivity of 92.5%, specificity of 94.6% and an AUC-ROC of 0.94 (Shahi et al., 2021).

Screening and definitive diagnosis of child abuse is by observation and analysis of the medical history, physical examination, laboratory tests and imaging study results. The increased accuracy rate of AI in the screening of child abuse could allow for earlier and immediate intervention of child abuse such as calling Child Protective Services (CPS).

AI Accuracy in Disease Diagnosis

Aside from screening for disease, AI can be used to improve the accuracy in disease diagnosis. There have been a number of clinical studies looking at the accuracy of AI in medical decision-making specifically pertaining to disease diagnosis.

Ovarian Cancer

The American Cancer Society states ovarian cancer accounts for more deaths than any other cancer of the female reproductive system in the U.S. and it estimates that about 13,770 women in the U.S. will die from ovarian cancer this year (Society, 2021d). Some risk factors for ovarian cancer include age over 40 years, obesity, smoking, women having their first full-term pregnancy after the age of 35 or never having a full-term pregnancy. Ovarian cancer diagnosis involves a pelvic exam done by a gynecologist (a doctor who specializes in female reproductive health), a blood test of a tumor marker called CA-125 and radiological exams such as a transvaginal ultrasound and CT scan of the pelvis (T. M. Clinic). A transvaginal ultrasound and CT scan of the pelvis are read and interpreted by a radiologist.

Akazawa and Hashimoto employed AI to predict ovarian cancer. They used the machine learning algorithm of XGBoost to diagnose ovarian cancer. They evaluated 202 women with ovarian tumors and used five machine learning classifiers including support vector machine, random forest, naive Bayes, logistic regression and XGBoost to obtain diagnostic results (Akazawa & Hashimoto, 2020) to diagnose ovarian cancer. They noted an 80% accuracy rate with the AI algorithm.

Definitive diagnosis of ovarian cancer is by an ovarian tissue biopsy, which is an invasive procedure. Stage 1 ovarian cancer is treated with surgery and depending on the grade, chemotherapy may also be required. Stage 2 and 3 ovarian cancer is treated with surgery and chemotherapy regardless of the grade. Lastly, stage 4 ovarian cancer is when the cancer has spread to different sites or organs in the body and it is very difficult to cure. Treatment involves surgery and chemotherapy (Society).

The increased accuracy rate of AI in the diagnosis of ovarian cancer could allow for detection and treatment at earlier stages of cancer and potentially increase the length of survival.

Pulmonary Hypertension

Pulmonary hypertension is a disorder in which the pressure in the blood vessels going from the heart to the lungs becomes high causing shortness of breath and/or chest pain. There is no cure for pulmonary hypertension and it progressively worsens over time. There are treatments available that can help alleviate symptoms and slow the progression of disease (T. M. Clinic). Thus, making it very crucial to detect pulmonary hypertension early so treatment can be initiated quickly. Initial treatment includes medications, however if medications do not alleviate symptoms then surgery is recommended.

Pulmonary hypertension diagnosis involves imaging tests such as a chest-x-ray and echocardiogram as well as an electrocardiogram (EKG). A chest x-ray is read by a radiologist, while an echocardiogram and EKG is read by a cardiologist, a medical doctor who specializes in the heart. Definitive diagnosis of pulmonary hypertension is by a right heart catheterization, which is an invasive procedure (T. M. Clinic). A right heart catherization is done by a cardiologist and it involves the placement of a thin, flexible catheter into the right side of the heart.

Kwon and colleagues developed an AI algorithm to identify pulmonary hypertension using an EKG. They evaluated 2939 individuals from two hospitals and noted that the AI algorithm demonstrated high accuracy in the detection of pulmonary hypertension (AUC-ROC 0.85).

Since there is no cure for pulmonary hypertension and it is a medical condition that progressively worsens over time, the initiation of medication earlier is critical. Medications available for pulmonary hypertension that can help alleviate symptoms and slow the progression of disease.

The increased accuracy rate of AI in the diagnosis of pulmonary hypertension could allow for early detection and early initiation of treatment. Subsequentially, it can improve the quality of life and potentially increase the length of survival as well.

Healthcare Associated Infections

The Centers for Disease Control and Prevention (CDC) defines infections called healthcare-associated infections (HAIs) as central line-associated bloodstream infection, catheter-associated urinary tract infection and ventilator-associated pneumonia (Prevention, 2014). Definitive diagnosis of a central line-associated bloodstream infection, catheter-associated urinary tract infection and ventilator-associated pneumonia is by either a blood, urine or sputum culture, respectively. HAIs contribute to a prolonged duration of hospital stay (Jia et al., 2019). In addition, some HAIs are potentially preventable complications of hospitalization for which hospitals will not receive payment by health insurance companies.

Tvardik and colleagues evaluated the performance of SYNODOS, an NLP solution for detecting medical events in electronic health records, in detecting healthcare-associated infections (Tvardik et al., 2018). The accuracy rate of automatic detection of healthcare-associated infections was 84%. The overall sensitivity was 83.9% and the specificity was 84.2%. Tvardik notes, “Automatic HAI detection algorithms could offer better surveillance standardization for hospital comparisons” (Tvardik et al., 2018).

The increased accuracy rate of AI in the diagnosis and detection of HAIs could help prevent hospitals from losing financial incentives for avoiding potentially preventable hospital complications such as HAIs.

Cervical and Anal Cancer

The American Cancer Society states that infection with the human papillomavirus (HPV) is the most common risk factor for cervical and anal cancer. In addition, women with a history of cervical cancer (or pre-cancer) have an increased risk of anal cancer (Society, 2020).

The diagnosis of cervical and anal HPV-associated cancer is dependent of the results of a histopathology result which is typically reviewed manually by a physician. Cervical cancer screening is done by a pap smear, a test in which a gynecologist scrapes and brushes cells from the cervix (Society). The histopathology from the cervical tissue cells is interpreted by a pathologist, a medical doctor who studies body fluids, tissues and organs. Anal cancer screening is done with a digital rectal exam. Definitive diagnosis of cervical cancer is by a cervical tissue biopsy and anal cancer is by an anal tissue biopsy, both of which are invasive procedures. If there is concern of anal cancer, then invasive tests such as an anoscope, proctoscope and/or anal ultrasound with an anal tissue biopsy is done (T. M. Clinic). The histopathology from the anal tissue cells is analyzed and interpreted by a pathologist.

Oliveria and colleagues at Yale University developed an Oliveria NLP algorithm for the identification of individuals with cancer or precancer of the cervix and anus from cervical and anal pathology reports (Oliveira et al., 2020). The NLP algorithm was developed, which incorporated machine learning and rule-based methods to extract diagnostic elements from the narrative pathology reports. The algorithm revealed over 90% accuracy in the identification of abnormal cytology, histology, and positive HPV tests.

Cervical cancer treatment is dependent on the stage and grade of disease as well the desire to maintain fertility. Stage 1, 2 and 3 cervical cancer typically involves a cone biopsy or surgery (dependent on the desire to maintain fertility) and/or radiation with or without chemotherapy dependent on the stage and grade of cancer. Stage 4 cervical cancer has spread outside the pelvis and it is no curable (Society). Anal cancer is treated with a combination of chemotherapy and radiation (T. M. Clinic).

The increased accuracy rate of AI in the diagnosis of cervical and anal cancer could allow for detection and treatment at earlier stages of cancer and potentially increase the length of survival.

Parkinson’s Disease

Parkinson’s disease is a neurodegenerative disease that affects movement and it is marked by tremor, slow gait and muscular rigidity (Aging). Nearly one million people are living with Parkinson’s disease in the U.S. and about 60,000 people are diagnosed annually (P. s. Foundation). AI is being utilized in neurology, a branch of medicine focused on the nervous system, in helping diagnosis Parkinson’s disease.

Belic and colleagues conducted a systematic review of 48 studies utilizing machine learning algorithms on data describing motions of upper and lower extremities in the diagnosis of Parkinson’s disease (Belic et al., 2019). The overall accuracy rate of AI diagnosis of Parkinson’s disease was about 90%.

Parkinson’s disease progresses slowly and the cause is unknown. It can be quite debilitating and an early diagnosis can permit treatment with medications sooner. Medications can be used to control symptoms and subsequently improve the quality of life. There are no specific tests to diagnose Parkinson’s disease (Aging). Physicians make a diagnosis based on physical signs and symptoms and a neurological physical exam.

The increased accuracy rate of AI in the diagnosis of Parkinson’s disease could permit for early detection and early initiation of treatment and a potential improvement in the quality of life with symptom control.

Gastric Cancer

Gastric cancer, also known as stomach cancer affects over 25,000 people living in the United States. Older age (age over 65 years) is a risk factor as well as the consumption of salted and smoked foods. The bacteria *H.plyori* is another risk factor for the development of gastric cancer (Society). According to the American Cancer Society, the relative 5-year survival rate for gastric cancer is 32%, making early diagnosis and treatment critical (Society, 2022). AI is being utilized in gastroenterology in helping diagnose gastric cancer using images from an endoscopy.

Hsiao and colleagues conducted a systematic review of eight studies utilizing machine learning algorithms on endoscopic images in the diagnosis of gastric cancer. (Hsiao et al., 2021). The overall accuracy rate of AI diagnosis of gastric cancer was over 80%.

Definitive diagnosis of gastric cancer is by an endoscopic biopsy, which is an invasive procedure. The increased accuracy rate of AI in the diagnosis gastric cancer could allow for detection and treatment at earlier stages of cancer and potentially increase the length of survival.

Serrated polyposis syndrome

Serrated polyposis syndrome, previously known as hyperplastic polyposis syndrome, is a rare condition characterized by serrated polyps in the colon. It is commonly seen in adults in their 50s and 60s. Diagnosis is done via a colonoscopy, which is a flexible tube with a camera at its end (C. Clinic). A colonoscopy is done by a gastroenterologist to examine the colon and rectum. Definitive diagnosis of serrated polyposis syndrome is by a biopsy of colon tissue, which is an invasive procedure (C. Clinic). The biopsy histopathology from the colonic tissue cells is interpreted by a pathologist.

Parthasarathy and colleagues used NLP on data from 323,494 colonoscopies and pathology reports and evaluated its accuracy in identifying patients with serrated polyposis syndrome. NLP outperformed manual review of biopsy results with an accuracy rate of 93%. Physicians diagnosed serrated polyposis syndrome in 38% of cases, however NLP was able to diagnose serrated polyposis syndrome two years earlier than the physician in 20% of these cases (Parthasarathy, Lopez, McMichael, & Burke, 2020).

Serrated polyposis syndrome is associated with colon cancer with 25-40% of people with serrated polyposis syndrome developing colorectal cancer (C. Clinic). The increased accuracy rate of AI in the diagnosis of serrated polyposis syndrome could allow for allow for detection and treatment at earlier stages of cancer and potentially increase the length of survival.

Kidney Stones

Kidney stones, also known as renal calculi or nephrolithiasis, are hard deposits of minerals in concentrated urine. Symptoms include back pain, nausea, vomiting, and if fever is present, there may be fever and chills (M. Clinic).

Diagnosis of kidney stones involves a urine test, blood test and radiological imaging such as CT scan of the abdomen or abdominal ultrasound. Kidney stones of small size are observed to see if they pass through the ureter and into the bladder on its own. Medications can be given to help expand the ureter to allow for easy passage of the kidney stone. A non-invasive treatment called lithotripsy can be done (T. M. Clinic). Lithotripsy is a shock wave treatment meant to break up kidney stones into smaller pieces allowing for easier passage of the stone through the ureters. If kidney stones are too large to pass through the ureters, then a more invasive procedure such as a percutaneous nephrolithotomy, may be required to remove the kidney stone.

Li and colleagues utilized NLP to identify CT Kidney, Ureter and Bladder (CT KUB) reports that were positive for kidney stones. They looked at a total of 1874 CT KUB reports. NLP revealed a 85% accuracy, 66% sensitivity and 95% specificity in the identification of kidney stones (Li & Elliot, 2019).

The increased accuracy rate of AI in the diagnosis of kidney stones could allow for early detection and treatment of kidney stone. In addition, it could avoid unnecessary and invasive procedures such as a percutaneous nephrolithotomy.

Alzheimer's Disease

Alzheimer's disease is type of dementia that is a progressive loss of memory and cognition. Early signs of Alzheimer's disease are memory impairment, trouble completing daily tasks and poor judgement in decisions. Diagnosis is through memory tests (mental status test and neuropsychological test) and radiological imaging such as a MRI or a CT scan of the head. Treatment includes medications that can help slow the progression of memory loss. However, there is no cure for Alzheimer's disease (M. Clinic).

Soni and colleagues used machine learning to predict the risk of Alzheimer's disease. Machine learning techniques used include random forest (RF), neural network (NN) and NLP. Machine learning had a higher accuracy rate of 76% than traditional methods of detecting Alzheimer's disease such as memory tests (mental status test and neuropsychological test) and radiological imaging (Soni, Amrhein, Baucum, Paek, & Khojandi, 2021).

The increased accuracy rate of AI in the diagnosis of Alzheimer's disease could allow for early detection and early initiation of treatment.

AI Accuracy in Predicting Survival Prognosis

Aside from diagnosing disease, AI can be used to improve the accuracy in predicting survival prognosis. There have been a number of clinical studies looking at the accuracy of AI in medical decision-making specifically pertaining to the prediction of survival.

Nasopharyngeal Cancer

Nasopharyngeal cancer is a cancer of the head and neck that involves the upper part of the throat that lies behind the nose called the pharynx. Initial diagnostic testing involves an invasive test called a nasal endoscopy. A nasal endoscopy is a thin, flexible tube with a camera on the end, and it is placed into the nose and advanced to the back of the throat. Definitive diagnosis of nasopharyngeal cancer is by a nasopharyngeal tissue biopsy, which is an invasive procedure (Society). The histopathology from the nasopharyngeal tissue biopsy is typically interpreted by a pathologist. The biopsy results along with results of the physical exam, imaging tests and blood tests are used by the oncologist, a medical doctor specializing in cancer, to determine the survival prognosis.

Akcay and colleagues utilized machine learning to evaluate the survival prognosis of nasopharyngeal cancer (Akcay, Etiz, Celik, & Ozen, 2020). There were 72 patients diagnosed with nasopharyngeal cancer. Machine learning algorithms used include logistic regression, artificial neural network, XGBoost, support-vector clustering, random forest, and Gaussian Naive Bayes, which resulted in a survival prognosis prediction accuracy of 77%, 88%, 77%, 33%, 66% and 88% respectively.

The increased accuracy rate of AI in determining the survival prognosis of nasopharyngeal cancer can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

Liver Cancer

The American Cancer Society reports that the incidence rates of liver cancer have more than tripled since 1980 and the death rates of liver cancer have more than doubled as well (Society, 2021b). Liver cancer is diagnosed by blood tests and imaging tests such as a CT scan, MRI or ultrasound. The imaging test results are read and interpreted by a radiologist. Definitive diagnosis of liver cancer is by a liver tissue biopsy, which is an invasive procedure. The histopathology of the liver tissue biopsy is read by the pathologist.

Quirino and colleagues conducted a literature review on studies done assessing the role of AI in the prediction of survival following hepatocellular carcinoma (HCC) treatment (Lai et al., 2020). Nine articles were identified with a patient sample size of 22,926. AI methodologies used include artificial neural networks (ANN), support vector machine, artificial plant optimization, and peritumoral radiomics. ANN was able to predict HCC survival with an 89.5% accuracy.

The increased accuracy rate of AI in determining the survival prognosis of liver cancer can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

Breast Cancer

The American Cancer Society states that breast cancer is the second leading cause of cancer death in women (Society, 2021a). Breast cancer has a high mortality rate making early diagnosis of breast cancer crucial. The diagnostic screening test for breast cancer is a mammogram, which is read by a radiologist (Force). Definitive diagnosis of breast cancer is by a breast tissue biopsy, which is an invasive procedure. The histopathology from the breast tissue biopsy is interpreted by a pathologist.

Montazeri and colleagues evaluated 900 patients and used machine learning techniques such as Naive Bayes (NB), Trees Random Forest (TRF), 1-Nearest Neighbor (1NN), AdaBoost (AD), Support Vector Machine (SVM), RBF Network (RBFN), and Multilayer Perceptron (MLP) for the prediction of breast cancer survival. The TRF technique showed the best results with the accuracy, sensitivity and AUC-ROC as 96%, 96% and 0.93, respectively (Montazeri, Montazeri, Montazeri, & Beigzadeh, 2016).

The increased accuracy rate of AI in determining the survival prognosis of breast cancer can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

Ovarian Cancer

According to the National Cancer Ovarian Coalition, ovarian cancer is the fifth most common cancer in women and if diagnosed and treated in the early stage, ovarian cancer has a 5-year survival rate over 90% (Coalition). Some risk factors for ovarian cancer include age over 40 years, obesity, smoking, women having their first full-term pregnancy after the age of 35 or never having a full-term pregnancy (Society). Ovarian cancer diagnosis involves a pelvic exam done by a gynecologist, a blood test of a tumor marker called CA-125 and radiological exams such as a transvaginal ultrasound and CT scan (T. M. Clinic).

Enshaei and colleagues evaluated 668 patients with ovarian cancer using AI methodologies such as artificial neural networks (ANN) to predict the survival prognosis. The model was able to predict ovarian cancer survival prognosis with an AUC of 0.73 and an accuracy rate of 77% (Enshaei, Robson, & Edmondson, 2015).

The increased accuracy rate of AI in determining the survival prognosis of ovarian cancer can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

Lung Cancer

According to the American Lung Association, the 5-year survival rate for lung cancer is 18.6% and only 16 percent of lung cancer cases are diagnosed at an early stage (A. L. Association). The most common symptoms of lung cancer are a persistent cough, vocal hoarseness, shortness of breath and unexplained weight loss (Society).

Wang and colleagues used machine learning algorithms to predict survival prognosis in 28,458 patients with primary lung cancer. In addition, they broke down their evaluation to males and females with primary lung cancer. The accuracy rate in predicting survival prognosis in males was 73.3% and in females it was 82.9% (Y. Wang et al., 2021).

The increased accuracy rate of AI in determining the survival prognosis of primary lung cancer can help improve patient outcomes with earlier detection, especially considering that only 16 percent of lung cancer cases are diagnosed at an early stage. In addition, it can permit providers to administer the best treatments quickly.

Acute Myeloid Leukemia

Acute Myeloid Leukemia (AML) is a cancer of the blood cells. Some risk factors include male sex, age over 65 years, smoking and exposure to chemicals such as benzene or formaldehyde (Society). There are about 20,000 new cases of AML annually and about 11,500 deaths each year from AML (Society).

Karami and colleagues predicted survival prognosis in 249 patients with AML using machine learning algorithms (Karami, Akbari, Moradi, Soleymani, & Fallahi, 2021). The AUC of AI for the prediction of the survival prognosis of AML was 0.93 and the accuracy rate reached was 85.2%.

The increased accuracy rate of AI in determining the survival prognosis of AML can help improve patient outcomes and allow providers to administer the best treatments more quickly.

Pancreatic Cancer

Pancreatic cancer is a cancer of the pancreas and it does not present in the early stage of the cancer because it rarely causes symptoms and goes unnoticed. Pancreatic cancer frequently presents at later stages making the survival prognosis quite low (T. M. Clinic). Since the year 2000, the incidence of pancreatic cancer has gone up one percent annually ("Pancreatic Cancer: Statistics,").

Pancreatic cancer typically presents with mid-abdominal pain that occasionally radiates to the back. It is associated with a loss of appetite, yellowing of the skin and eyes as well as weight loss. Diagnosis is through medical history, physical examination and radiological imaging such as CT scan, MRI or ultrasound of the abdomen. A specialized test called an endoscopic retrograde cholangiopancreatography (ERCP) is used to visualize the pancreatic duct and bile ducts (T. M. Clinic). An ECRP is a small thin tube with a camera on its end is passed down the mouth through the stomach and into the first part of the small intestine. Definitive diagnosis of pancreatic cancer is by a pancreatic tissue biopsy, which is an invasive procedure.

Bakasa and colleagues identified a total of 690 patients (370 patients with pancreatic ductal adenocarcinoma and 320 controls). ML algorithms were able to predict pancreatic ductal adenocarcinoma survival with an 98.8% accuracy, 99.3% specificity and 98.3% sensitivity (Bakasa & Viriri, 2021).

The increased accuracy rate of AI in determining the survival prognosis of pancreatic cancer can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

Glioblastoma

Glioblastoma is an aggressive type of brain cancer that is fast-growing. Symptoms depend on the location of the brain tumor and it includes vomiting, persistent headache, double or blurry vision, new-onset seizures and changes in mood or the ability to think (T. M. Clinic).

Diagnosis is via CT scan or MRI of the brain. The Response Assessment in Neuro-Oncology (RANO) is the traditional method of assessing the MRI scan of glioblastoma (Chukwueke & Wen, 2019). However, this assessment is inaccurate and unreliable since it relies on manual two-dimensional measurements of contrast-enhancing target lesions.

Kickingereder and colleagues utilized machine learning algorithms that used artificial neural networks (ANN) to create a 3-plan analysis of the MRI scans. They evaluated MRI scans of 455 patients with brain tumors and found ANN to be 36% more effective than traditional RANO methods of assessment and in predicting overall survival (Kickingereder et al., 2019).

The increased accuracy rate of AI in determining the survival prognosis of glioblastoma can help improve patient outcomes by allowing providers to administer the best treatments more quickly.

AI Accuracy in Predicting Disease Complication

Aside from utilizing AI to predict survival prognosis, AI can be utilized to predict disease complication. There have been a number of clinical studies looking at the accuracy of AI in medical decision-making specifically pertaining to the prediction of the development of disease complication.

Diabetic Kidney Disease

Diabetic kidney disease, also known as diabetic nephropathy, is the leading cause of kidney failure and over 247,000 people are living with diabetic kidney disease in the United States according to the National Kidney Foundation (N. K. Foundation).

Diabetic kidney disease is typically screened for during annual physical exam visits with a blood glucose test, blood basic metabolic panel test, urinalysis and urinary albumin test. Additional diagnostic tests include radiological imaging tests such as an x-ray or ultrasound and a kidney biopsy. Treatment includes medication along with a diabetic diet and exercise. If diabetic kidney disease is advanced, kidney dialysis or a kidney transplant may be needed (T. M. Clinic). Dialysis has a significant impact on one’s quality of life considering dialysis typically requires a visit to the dialysis center 3 times a week with each session lasting 3 to 5 hours.

Makino and colleagues utilized machine learning algorithms on the medical records of 64,059 diabetic patients to predict diabetic kidney disease (Makino et al., 2019). AI was able to predict diabetic kidney disease in these patients with 71% accuracy and AUC of 0.743.

The increased accuracy rate of AI in predicting diabetic kidney disease in diabetics can help improve patient outcomes by allowing providers to administer the best treatments more quickly. Additionally, it improves the quality of life by avoiding unnecessary and invasive procedures such as a kidney biopsy, dialysis and kidney transplant.

Asthma Exacerbation

Asthma is a result of swelling and narrowing of the airways that carry air from the nose and mouth to the lungs. According to the Asthma and Allergy Foundation of America, there are approximately 25 million people living with asthma in the United States and on average, ten people die from asthma daily (America).

Symptoms include shortness of breath, chest tightness, wheezing and coughing Asthma exacerbations can be potentially life-threatening if not treated in a timely manner. Treatment involves medications, however if an asthma exacerbation progressively worsens, then a more invasive procedure called intubation may be needed to provide air to the lungs (T. M. Clinic).

Finkelstein and Wood used machine learning algorithms to predict asthma exacerbations on 26 adult patients. The accuracy rate of the AI prediction of asthma exacerbation was 80%. The sensitivity and specificity of AI prediction of asthma exacerbation was 84% and 80% respectively (Finkelstein & Wood, 2013).

The increased accuracy rate of AI in predicting asthma exacerbation can help improve patient outcomes by allowing providers to administer the asthma treatments more quickly. Additionally, it could avoid unnecessary and invasive procedures such as airway intubation.

Age- Related Macular Degeneration

Age-related macular degeneration is loss of central vision and it is the leading cause of vison loss in individuals over the age of 50 years (Medicine).

Common signs and symptoms include blurry vision, reduced central vision or difficulty recognizing faces. Diagnosis of age-related macular degeneration is by a dilated eye exam or an optical coherence tomography which is done by an ophthalmologist. Additionally testing with a fluorescein angiography may be needed, which is a more invasive procedure. Treatment can be with medications or involve invasive treatments such as laser therapy (Opthalamology).

Bhuiyan and colleagues built and validated machine learning algorithms to predict age-related macular degeneration progression in 4139 patients. The accuracy rate for AI prediction of age-related macular degeneration progression was noted to be 84% (Bhuiyan et al., 2020).

The increased accuracy rate of AI in predicting age-related macular degeneration progression could permit early intervention with medication and a potential improvement in the quality of life.

Non-ST-Elevation Myocardial Infarction (NSTEMI)

Non-ST-Elevation Myocardial Infarction (NSTEMI) is a type of acute coronary syndrome (ACS), also known as a heart attack. ACS is associated with a high morbidity and mortality and prompt diagnosis is critical. Common symptoms include chest tightness, shortness of breath, nausea and indigestion. Diagnosis is made by a blood test and an EKG (T. C. Clinic). Treatment includes medications and if there is advanced coronary artery disease, then cardiac catheterization, an invasive procedure, may be needed.

Wu and colleagues utilized machine learning algorithms to evaluate 268 patients with chest pain to predict NSTEMI. The machine learning methodology used was artificial neural networks (ANN). ANN was able to predict ACS in patients with chest pain with a sensitivity of 90.9%, a specificity of 93.3% and an accuracy of 92.9% (Wu et al., 2019).

The increased accuracy rate of AI in predicting NSTEMI in patients with chest pain could permit early intervention with medication and the avoidance of invasive procedures such as cardiac catheterization.

Critical COVID-19 Infection

Coronavirus disease (COVID-19) is an infectious disease cause by the SARS-CoV-2 virus that was responsible for the COVID-19 pandemic. COVID-19 is spread by person to person from respiratory droplets. Symptoms include cough, fever, chills, sore throat, loss of taste or smell, and shortness of breath. COVID-19 can cause mild, moderate and severe (critical) illness (M. Clinic).

The Acute Physiology and Chronic Health Evaluation (APACHE II) score estimates ICU mortality based on current physiologic measurements, age and previous health conditions (Luo, Wang, & Wang, 2021). The APACHE score is the currently used clinical status severity score and mortality estimation tool.

Assaf and colleagues utilized machine learning models to predict the risk of severe COVID-19 illness based on their clinical status on hospital admission. They evaluated 6995 patients, out of which 162 were admitted for non-severe COVID infection. Of the 162 patients, 25 (15.4%) patients developed critical COVID-19 (Assaf et al., 2020). Machine learning models used include artificial neural network (ANN), random forest (RF) and classification and regression tree (CRT). The machine learning models had higher sensitivity (88.0%), specificity (92.7%) and accuracy (92.0%) than the currently used APACHE II score in predicting the risk of severe COVID-19 infection (Assaf et al., 2020).

The increased accuracy rate of AI in predicting critical COVID-19 illness in patients admitted for mild COVID infection could permit early intervention with medication and the avoidance of admission into the ICU.

Prematurity, Preeclampsia and Neonatal Mortality in Pregnancy

According to the World Health Organization, fifteen percent of pregnant women are at risk of developing a life-threatening complication and 800 women died worldwide annually from complications developed from pregnancy (Organization). Some common potential complications associated with pregnancy include prematurity, preeclampsia, high blood pressure and gestational diabetes (Health). Prematurity of birth is when a baby is born too early, that is before 37 weeks of gestation (T. M. Clinic). Preeclampsia is when a pregnant women develops high blood pressure and signs of damage to organs such as the liver and kidney (T. M. Clinic). Prematurity and preeclampsia are serious complications of pregnancy that should be prevented.

Bertini and colleagues conducted a systematic review of 31 studies utilizing machine learning algorithms to predict the risk of developing complications of pregnancy (Bertini, Salas, Chabert, Sobrevia, & Pardo, 2021). Data used by machine learning included electronic medical records, medical images, biological markers and fetal heart rate sensors. The machine learning algorithms outperformed in the prediction of prematurity, preeclampsia and neonatal mortality.

The machine learning technique called support vector machine was used to predict prematurity from medical images with an accuracy of 95.7% and an AUC of 0.952. The machine learning technique AdaBoost model was utilized to predict preeclampsia with an accuracy of 89% and an AUC of 0.964. The machine learning technique XGBoost was used to predict neonatal mortality from socioeconomic demographics with an accuracy rate of 99.7% and an AUC of 0.842 (Bertini et al., 2021).

The increased accuracy rate of AI in predicting pregnancy complications in pregnant women could permit early intervention, the avoidance of preterm labor and potentially neonatal death.

Stroke Risk with Atrial Fibrillation

Atrial fibrillation is a very rapid and irregular heart rhythm (arrythmia) that can lead to the development of blood clots. These blood clots can travel to the heart and cause a heart attack or it can travel to the head and cause a stroke. Both a heart attack and stroke can be potentially fatal (T. M. Clinic). Atrial fibrillation is the most common heart arrythmia and it is a risk factor for the development of a stroke. Atrial fibrillation is associated with a five-fold increase of ischemic stroke (Virani et al., 2021).

Risk factors include advanced age (over 65 years), obesity, high blood pressure, heart failure, ischemic heart disease and diabetes. Atrial fibrillation is diagnosed mainly with an EKG and treated with medication to control the heart rhythm, blood-thinning medication to prevent blood clots and/or surgery (T. M. Clinic).

Wang and colleagues reviewed 480 electronic medical records of patients with atrial fibrillation. They tested NLP algorithms to predict the likelihood of stroke. Although they did not evaluate the accuracy of the NLP algorithms, they calculated the positive predictive value (PPV) (S. V. Wang, Rogers, Jin, Bates, & Fischer, 2017). The PPV, also known as precision, tells you how likely someone who tests positive actually has the disease (Safari et al., 2015). Wang and colleagues noted they were able to better identify and target patients at high risk for stroke.

The increased accuracy rate of AI in predicting the potential of stroke in patients with atrial fibrillation could permit early intervention with medication and potentially improve the quality of life.

Outcomes in Acute Ischemic Stroke

An ischemic stroke is when blood supply to a part of the brain is interrupted and prevents brain tissue from receiving oxygen and nutrients. It is considered a medical emergency and immediate medical attention is required. Symptoms of stroke include headache, numbness and trouble speaking, walking or seeing (A. S. Association). According to the CDC, one in every six death from cardiovascular disease is from a stroke and 87% of strokes are ischemic strokes (C. f. D. C. a. Prevention).

The Acute Stroke Registry and Analysis of Lausanne (ASTRAL) score is score based on epidemiological, clinical, laboratory and brain imaging data that is used to predict the functional outcome after an ischemic stroke. Heo and colleagues developed three machine learning models and compared its accuracy to the ASTRAL score in predicting long-term outcomes in ischemic stroke. They examined 2604 patients in total. The machine learning algorithms outperformed the ASTRAL score in predicting long-term outcomes. Of the 2604 patients, 2043 (78%) of them had favorable outcomes (Heo et al., 2019).

The increased accuracy rate of AI in predicting long-term outcomes in ischemic stroke patients could permit early intervention with medication and a potential improvement in the quality of life.

Subdural Hematoma

Subdural hematoma is an accumulation of blood on the surface of the brain. It is considered a medical emergency. Symptoms include persistent headache, slurred speech, vision changes, confusion, nausea and vomiting. Diagnosis is by medical history and radiological imaging such a CT scan of the head. Subdural hematomas that cause mild or no symptoms can be observed with bed rest and medication, while larger hematomas require decompression brain surgery (T. C. Clinic).

Pruitt and colleagues evaluated 643 CT scan of heads of patients who sustained subdural hematoma. They utilized NLP algorithms to predict complications of subdural hematoma. NLP outperformed with higher accuracy of 92% in predicting complications of subdural hematoma.

The increased accuracy rate of AI in predicting complications of subdural hematoma could permit early intervention with medication and a potential improvement in the quality of life.

Chapter IV.

Discussion

AI Accuracy in Healthcare

AI is becoming increasingly integrated into healthcare and it has the potential revolutionize the delivery of medical care. A number of research studies have shown AI algorithms to outperform clinicians in screening and detecting disease, as well in predicting survival prognosis and disease complication. These studies reveal higher accuracy rates among AI systems, with accuracy rates above 75%. AI has the potential to limit variation and improve accuracy in clinical decisions, as well prevent avoidable medical errors. Consequently, AI has shown to increase productivity, improve operational efficiency and reduce medical cost.

According to a report by the consulting firm Accenture, the application of AI in the healthcare sector can create $150 billion in annual savings by 2026 (Accenture). The use of AI in hospitals has resulted in improved precision and accuracy in medical diagnosis and it appears to be worth the investment to utilize AI in the healthcare setting. Although research studies on AI in healthcare demonstrate increased accuracy and precision in clinical diagnosis, AI comes with some functional and ethical challenges.

Functional Challenges of AI in Healthcare

Functional challenges in the implementation of AI in the healthcare setting include large datasets of patient information that are siloed and non-standardized among various healthcare entities, the lack of transparency and sharing of data among these stakeholders, the negative impact of AI integration on a provider's clinical workflow and the need for quality control.

Siloed, Non-Standardized and Large Datasets

Healthcare has its own set of unique functional challenges. Unlike other industries that utilize AI, the healthcare industry's utilization of AI is limited due to very large datasets that are siloed and not standardized (McQuitty, 2020). Healthcare data is mainly held by hospitals, health insurance companies and pharmaceutical companies. Healthcare data sets in inpatient settings include information on patient demographics, diagnoses, medical history, surgeries, medications, laboratory tests, radiological imaging and health insurance data. Health insurance company datasets include information on demographics, billing claims, physician visits, diagnoses and pharmacy data. Pharmaceutical companies have detailed information on patient medication history and pharmacy utilization data. Aside from patient data held in electronic health records, patient data is also available from wearable devices, smartphone and mobile apps. Unfortunately, each of these healthcare sectors that hold large volumes of patient data are not easily shared.

Moreover, patients switch providers and health insurances as well as go to different hospitals which lead to data being split among different stakeholders and saved in multiple formats. These factors prevent interoperability and consequentially compromise patient safety and the quality of medical care (Heath, 2016).

In addition, healthcare data is not standardized and it is held by various stakeholders (hospital, provider and health insurer) in different electronic formats (Management). Each stakeholder collects and owns specific types of patient data that the other stakeholder needs and is sometimes unwilling to share.

Lack of Transparency and Poor Collaboration

The lack of transparency and poor collaboration among stakeholders creates an obstacle in integrating AI in healthcare. Gerke and colleagues state “transparency creates trust among stakeholders, particularly clinicians and patients, which is the key to successful implementation of AI in clinical practice” (Gerke, Minssen, & Cohen, 2020). Medical decisions on treatment and diagnoses are dependent on the access to timely and complete patient information. The usefulness of health data is dependent on the quality and completeness of the data. In order for AI to work effectively, the data inputs it receives must be complete and accurate. Incomplete health data contributes to an increased likelihood of redundant testing, transcription errors and adverse events. Bowman states “ Poor EHR system design and improper use can cause EHR-related errors that jeopardize the integrity of the information in the EHR, leading to errors that endanger patient safety or decrease the quality of care (Bowman, 2013).

Workflow Integration

The workflow processes providers follow to care for patients are not standardized and when these workflows are redesigned to allow for interoperability, it causes a significant negative impact on a provider’s day-to-day workflow. Medical staff have to learn to incorporate AI-enabled real-time insights into their existing workflow (CIO). Providers should not be inundated with an abundance of clinical alerts. If the integration of AI adds additional steps in the clinical workflow, it will add more strain on already time constrained schedule, and decrease productivity and operational efficiency.

In addition to the change in a provider’s clinical workflow, there are also constraints on a provider’s time for the required education and training on the AI system.

The key is finding a balance in the frequency of AI alerts to the provider as well as developing a seamless integration into the providers' clinical workflow.

Data Gathering and Quality Control

AI systems require a large amount of data and it is crucial that the data an AI system is presented come from reliable sources. Data from unreliable sources can affect the AI output negatively and have the system make incorrect decisions or suggestions. What if the AI system fails to detect a tumor on an MRI scan or it recommends the wrong medication?

AI errors are different from human errors for two main reasons. First, a diagnostic error from an AI system has the potential to affect a large number of patients rather than a single patient with a human diagnostic error. Second, patients may react to an error made from a AI system differently than to an error made by their physician (Report). Patients generally believe and trust that their physician are acting towards them with goodwill while that same sentiment may not exist when an AI system is making a clinical decision about them.

A regulatory framework must be in place to ensure rigorous quality control to maintain safe and effective AI algorithms. Without a regulatory framework and quality controls in place, significant risks to patient care exist. Periodic system-wide updates and improvements along with ongoing performance monitoring and provider feedback need to be used to continually adjust and regulate the AI system. In addition, when risks present, they should be corrected immediately.

Ethical Challenges of AI in Healthcare

Ethical challenges of AI implementation in healthcare include the ownership of patient data, patient autonomy, accountability for negative consequences, the risk for biases and discrimination and the threat to data privacy and security.

Ownership

Ethical challenges surrounding AI in healthcare include the ownership of patient information, risk of bias and the threat to data privacy and security. With regards to ownership of patient health data, questions arise such as should patients have ownership over their data and if not, should they be made aware or consent before their medical information is shared? If patients are not given ownership to their health data, should they have a right to decide how their medical data is used?

Precluding patients from ownership of their health data may cause patients not to disclose sensitive personal information which may result in an incomplete health profile being utilized to make a medical decision. Incomplete health data can result in potential errors and adverse events. Patients may be weary to provide sensitive personal information, such as sexual health information or drug use, fearing it may be used against them by health insurance companies or they may be anxious over the unknowingness of who has access to their data.

Patients also worry that payers will use negative health data (chronic medical conditions or smoking habit) to increase premiums on their health insurance plan (Times). Additionally, patients are concerned over the privacy of their data since they are unaware of all who have access to their personal health data.

Patient autonomy

However, providing patients with complete autonomy over their health data could result in patients omitting vital medical information that they do not want to share. When patients gain access to their medical records, they gain insight into their medical conditions and may be tempted to remove embarrassing or sensitive information related to their medical history resulting in an incomplete health profile. Patient may want to remove sensitive information surrounding their sexual health or substance abuse that they had previously disclosed. Incomplete or inaccurate health data increases the risk of potential errors and adverse events.

Accountability

As AI becomes increasingly utilized in the healthcare setting to make medical decisions pertaining to disease screening and detection, as well in predicting survival prognosis and disease complication, the question arises as to who is responsible for negative consequences secondary to an AI decision. For example, what if the AI algorithm misses a cancer diagnosis or it predicts a lower disease survival prognosis than what it truly is? Who should take accountability for the misdiagnosis or misinterpretation? Would it be the AI system developers, the medical staff or the hospital? (Lee & Yoon, 2021). The challenge remains in determining who has oversight and accountability on the final clinical determination.

Biases and Discrimination

AI also bears a risk for biases and discrimination that can result in harmful patient outcomes. Unrepresentative samples, whether it is due to race, gender, age or health conditions, results in inaccurate decision-making by AI. Panch and colleagues explain bias in the context of AI as “the instances when the application of an algorithm compounds existing inequities in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and adversely impact inequities in health systems” (Panch, Mattie, & Atun, 2019). The inequitable collection of data leads to lack of inclusion and subsequentially the AI algorithm prediction to be biased.

AI algorithms are based on the data it is presented. According to Keskinbora, “the AI algorithms written could naturally contain errors that may result in unforeseen consequences and unfair outcomes along economic and racial class lines” (Keskinbora, 2019).

To mitigate biases, the data used to train AI must be representatively diverse as well as in the people collecting the data. Large data sets and diversity in the training of the algorithms is essential in avoiding generalization. Datasets should be representative of a demographically diverse population and be of high quality in order to be useful. Kaushal and colleagues explain, “because obtaining high-quality data is challenging, researchers are also building algorithms that try to do more with less. But for now, ensuring diversity of data used to train algorithms is central to our ability to understand and mitigate biases of AI” (Kaushal, 2020).

Additionally, “black box” algorithms used in AI generate some ethical and liability concerns when the decision-making process from when the data is fed into the model to when the output is delivered is not fully understood. The inputs and outputs are known, however the process of the “black box” system is hidden and unknown. When an undesired outcome occurs and there is a need to troubleshoot, it becomes difficult to find the exact cause because the process is not known. Bresnick says, “black-box artificial intelligence tools that give little rationale for their decisions only complicate the problem – and make it more difficult to assign responsibility to an individual when something goes awry” (Bresnick, 2018a). Algorithmic bias and “black box” algorithms must be addressed to ensure successful integration of AI in healthcare.

Data Privacy and Security

AI also poses a threat to data privacy and security. Large quantities of patient data are in the hands of multiple entities (hospital, physician/provider, health insurance) which can add to privacy risks and concerns. The multiplicity of parties handling data also compromises the security of the heath information. Sulmasy and colleagues state, “breeches may occur accidently, through cyber-attacks, or due to lapses in professional conduct, such as searching for test results of a family member” (Sulmasy, Lopez, Horwitch, Professionalism, & Human Rights, 2017).

To protect data integrity, policies and procedures to maintain patient privacy and security need to be put in place. Patient data should be anonymized and deidentified. In addition, security measures such as encryption, antivirus software, firewalls and routine random audits should be established to prevent the inappropriate use of datasets. Blockchain technology could potentially play a role in the future as a platform for secure digital health information exchange among multiple stakeholders.

AI in Healthcare

AI-augmented services are programmed to provide benefits to its’ stakeholders; however, it can also present some obstacles and risks. Healthcare is an industry that possesses distinctive ethical challenges that must be taken into consideration when implementing AI. The ownership of patient information, risk of bias, threat to data privacy and security and the lack of interoperability in healthcare compromises patient safety and the quality of medical care. With the implementation of any new technology in healthcare, there will arise ethical challenges that need to be evaluated and addressed so that patient safety and the quality of medical care are not compromised.

Awareness of these challenges and working within an ethical framework should guide the development of AI systems by developers and designers. The collection, sharing and storage of patient data should be done in accordance to privacy laws. In addition, data should only be accessible by authorized users and when data is transmitted it should always be encrypted. Data minimization should be incorporated, meaning that only necessary patient data should be utilized and any data no longer needed should be deleted. Lastly, AI systems should regularly be monitored and when risks present they should be corrected immediately.

Many studies showed improved precision and accuracy in medical decision-making with the utilization of AI. Studies evaluated the accuracy of AI in medical decision-making pertaining to disease screening, diagnosis of disease and predicting survival prognosis and disease complication. Although, there are some potential risks and challenges associated with the utilization of AI in the healthcare setting, the benefits appear to outweigh the risks. AI is an evolving and growing field in healthcare. Further studies need to be conducted to reinforce its accuracy and precision in clinical decision-making.

References

Accenture. Artifical Intelligence: Healthcare's New Nervous System Retrieved from https://www.accenture.com/_acnmedia/pdf-49/accenture-health-artificial-intelligence.pdf

Aging, T. N. I. o. Parkinson's Disease. Retrieved from https://www.nia.nih.gov/health/parkinsons-disease

Akazawa, M., & Hashimoto, K. (2020). Artificial Intelligence in Ovarian Cancer Diagnosis. *Anticancer Res, 40*(8), 4795-4800. doi:10.21873/anticanres.14482

Akcay, M., Etiz, D., Celik, O., & Ozen, A. (2020). Evaluation of Prognosis in Nasopharyngeal Cancer Using Machine Learning. *Technol Cancer Res Treat, 19*, 1533033820909829. doi:10.1177/1533033820909829

Alliance, T. N. C. s. National Statistics On Child Abuse. Retrieved from https://www.nationalchildrensalliance.org/media-room/national-statistics-on-child-abuse/

America, A. a. A. F. o. Asthma Facts and Figures. Retrieved from https://www.aafa.org/asthma-facts/

Assaf, D., Gutman, Y., Neuman, Y., Segal, G., Amit, S., Gefen-Halevi, S., . . . Tirosh, A. (2020). Utilization of machine-learning models to accurately predict the risk for critical COVID-19. *Intern Emerg Med, 15*(8), 1435-1443. doi:10.1007/s11739-020-02475-0

Association, A. D. (2021). Statistics About Diabetes. Retrieved from https://www.diabetes.org/resources/statistics/statistics-about-diabetes

Association, A. L. Lung Cancer Fact Sheet. Retrieved from https://www.lung.org/lung-health-diseases/lung-disease-lookup/lung-cancer/resource-library/lung-cancer-fact-sheet

Association, A. S. Ischemic Stroke. Retrieved from https://www.stroke.org/en/about-stroke/types-of-stroke/ischemic-stroke-clots

Bakasa, W., & Viriri, S. (2021). Pancreatic Cancer Survival Prediction: A Survey of the State-of-the-Art. *Comput Math Methods Med, 2021*, 1188414. doi:10.1155/2021/1188414

Bang, C. S., Lee, J. J., & Baik, G. H. (2020). Artificial Intelligence for the Prediction of Helicobacter Pylori Infection in Endoscopic Images: Systematic Review and Meta-Analysis Of Diagnostic Test Accuracy. *J Med Internet Res, 22*(9), e21983. doi:10.2196/21983

Baratloo, A., Hosseini, M., Negida, A., & El Ashal, G. (2015). Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. *Emerg (Tehran), 3*(2), 48-49. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26495380

Belic, M., Bobic, V., Badza, M., Solaja, N., Duric-Jovicic, M., & Kostic, V. S. (2019). Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease-A review. *Clin Neurol Neurosurg, 184*, 105442. doi:10.1016/j.clineuro.2019.105442

Bertini, A., Salas, R., Chabert, S., Sobrevia, L., & Pardo, F. (2021). Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review. *Front Bioeng Biotechnol, 9*, 780389. doi:10.3389/fbioe.2021.780389

Bhuiyan, A., Wong, T. Y., Ting, D. S. W., Govindaiah, A., Souied, E. H., & Smith, R. T. (2020). Artificial Intelligence to Stratify Severity of Age-Related Macular Degeneration (AMD) and Predict Risk of Progression to Late AMD. *Transl Vis Sci Technol, 9*(2), 25. doi:10.1167/tvst.9.2.25

Bowman, S. (2013). Impact of electronic health record systems on information integrity: quality and safety implications. *Perspect Health Inf Manag, 10*, 1c. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24159271

Bresnick, J. (2016). What Is the Role of Natural Language Processing in Healthcare? Retrieved from https://healthitanalytics.com/features/what-is-the-role-of-natural-language-processing-in-healthcare

Bresnick, J. (2018a). Arguing the Pros and Cons of Artificial Intelligence in Healthcare. Retrieved from https://healthitanalytics.com/news/arguing-the-pros-and-cons-of-artificial-intelligence-in-healthcare

Bresnick, J. (2018b). What Is Deep Learning and How Will It Change Healthcare? Retrieved from https://healthitanalytics.com/features/what-is-deep-learning-and-how-will-it-change-healthcare

Campbell, J. (2015). Will we ever achieve interoperability? Not until there are incentives to do so. Retrieved from https://www.beckershospitalreview.com/healthcare-information-technology/will-we-ever-achieve-interoperability-not-until-there-are-incentives-to-do-so.html?oly_enc_id=2660C1094034B2W

Canchi, T., Ng, E. Y., Narayanan, S., & Finol, E. A. (2018). On the assessment of abdominal aortic aneurysm rupture risk in the Asian population based on geometric attributes. *Proc Inst Mech Eng H, 232*(9), 922-929. doi:10.1177/0954411918794724

Chowdhury, M. A., Apon, A., & Dey, K. (2017). *Data analytics for intelligent transportation systems*. Amsterdam: Elsevier.

Chukwueke, U. N., & Wen, P. Y. (2019). Use of the Response Assessment in Neuro-Oncology (RANO) criteria in clinical trials and clinical practice. *CNS Oncol, 8*(1), CNS28. doi:10.2217/cns-2018-0007

CIO. AI in healthcare: The tech is here, the users are not. Retrieved from https://www.cio.com/article/191781/ai-in-healthcare-the-tech-is-here-the-users-are-not.html

Clinic, C. Serrated Polyposis Syndrome/Hyperplastic Polyposis Syndrome. Retrieved from https://my.clevelandclinic.org/health/diseases/17462-serrated-polyposis-syndromehyperplastic-polyposis-syndrome

Clinic, M. Coronavirus disease 2019. Retrieved from https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963

Clinic, M. Diagnosing Alzheimer's: How Alzheimer's is diagnosed. Retrieved from https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers/art-20048075

Clinic, M. Kidney stones overview. Retrieved from https://www.mayoclinic.org/diseases-conditions/kidney-stones/symptoms-causes/syc-20355755

Clinic, T. C. Abdominal Aortic Aneurysm. Retrieved from https://my.clevelandclinic.org/health/diseases/7153-abdominal-aortic-aneurysm

Clinic, T. C. Diabetic Retinopathy. Retrieved from https://my.clevelandclinic.org/health/diseases/8591-diabetic-retinopathy

Clinic, T. C. H.pylori Infection. Retrieved from https://my.clevelandclinic.org/health/diseases/21463-h-pylori-infection

Clinic, T. C. NSTEMI: Non-ST-Elevation Myocardial Infarction (Heart Attack). Retrieved from https://my.clevelandclinic.org/health/diseases/22233-nstemi-heart-attack

Clinic, T. C. Subdural Hematoma. Retrieved from https://my.clevelandclinic.org/health/diseases/21183-subdural-hematoma

Clinic, T. M. Abdominal Aortic Aneurysm. Retrieved from https://www.mayoclinic.org/diseases-conditions/abdominal-aortic-aneurysm/symptoms-causes/syc-20350688

Clinic, T. M. Anal Cancer. Retrieved from https://www.mayoclinic.org/diseases-conditions/anal-cancer/symptoms-causes/syc-20354140

Clinic, T. M. Asthma. Retrieved from https://www.mayoclinic.org/diseases-conditions/asthma/symptoms-causes/syc-20369653#:~:text=Asthma%20is%20a%20condition%20in,asthma%20is%20a%20minor%20nuisance.

Clinic, T. M. Atrial Fibrillation. Retrieved from https://www.mayoclinic.org/diseases-conditions/atrial-fibrillation/symptoms-causes/syc-20350624#:~:text=Atrial%20fibrillation%20(A%2Dfib),and%20other%20heart%2Drelated%20complications.

Clinic, T. M. Diabetic nephropathy. Retrieved from https://www.mayoclinic.org/diseases-conditions/diabetic-nephropathy/symptoms-causes/syc-20354556

Clinic, T. M. Glioblastoma. Retrieved from https://www.mayoclinic.org/diseases-conditions/glioblastoma/cdc-20350148#:~:text=Glioblastoma%20is%20an%20aggressive%20type%20of%20cancer%20that%20can%20occur,%2C%20nausea%2C%20vomiting%20and%20seizures.

Clinic, T. M. Kidney Stones. Retrieved from https://www.mayoclinic.org/diseases-conditions/kidney-stones/symptoms-causes/syc-20355755

Clinic, T. M. Melanoma Symtpoms, Causes, Diagnosis and Treatment. Retrieved from https://www.mayoclinic.org/diseases-conditions/melanoma/diagnosis-treatment/drc-20374888

Clinic, T. M. Ovarian Cancer. Retrieved from https://www.mayoclinic.org/diseases-conditions/ovarian-cancer/diagnosis-treatment/drc-20375946

Clinic, T. M. Pancreatic Cancer. Retrieved from https://www.mayoclinic.org/diseases-conditions/pancreatic-cancer/symptoms-causes/syc-20355421#:~:text=Pancreatic%20cancer%20is%20cancer%20that,help%20manage%20your%20blood%20sugar.

Clinic, T. M. Preeclampsia. Retrieved from https://www.mayoclinic.org/diseases-conditions/preeclampsia/symptoms-causes/syc-20355745#:~:text=Preeclampsia%20is%20a%20pregnancy%20complication,blood%20pressure%20had%20been%20normal.

Clinic, T. M. Premature birth. Retrieved from https://www.mayoclinic.org/diseases-conditions/premature-birth/symptoms-causes/syc-20376730#:~:text=In%20other%20words%2C%20a%20premature,higher%20the%20risk%20of%20complications.

Clinic, T. M. Pulmonary Hypertension. Retrieved from https://www.mayoclinic.org/diseases-conditions/pulmonary-hypertension/symptoms-causes/syc-

20350697#:~:text=Pulmonary%20hypertension%20is%20a%20type,are%20narrowed%2C%20blocked%20or%20destroyed.

Coalition, N. O. C. General Prognosis. Retrieved from https://ovarian.org/about-ovarian-cancer/what-is-ovarian-cancer/?utm_term=ovarian%20cancer%20statistics&utm_campaign=About+Ovarian+Cancer+1&utm_source=adwords&utm_medium=ppc&hsa_acc=9835623983&hsa_cam=12621261328&hsa_grp=123687503967&hsa_ad=509726847358&hsa_src=g&hsa_tgt=kwd-93701016&hsa_kw=ovarian%20cancer%20statistics&hsa_mt=b&hsa_net=adwords&hsa_ver=3&gclid=EAIaIQobChMIp6e5hova9QIVGfjICh1bKwREEAAYASAAEgKnVPD_BwE

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthc J, 6*(2), 94-98. doi:10.7861/futurehosp.6-2-94

Enshaei, A., Robson, C. N., & Edmondson, R. J. (2015). Artificial Intelligence Systems as Prognostic and Predictive Tools in Ovarian Cancer. *Ann Surg Oncol, 22*(12), 3970-3975. doi:10.1245/s10434-015-4475-6

Finkelstein, J., & Wood, J. (2013). Predicting asthma exacerbations using artificial intelligence. *Stud Health Technol Inform, 190*, 56-58. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23823374

Force, U. S. P. S. T. Abdominal Aortic Aneurysm: Screening. Retrieved from https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/abdominal-aortic-aneurysm-screening

Force, U. S. P. S. T. Breast Cancer: Screening. Retrieved from https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/breast-cancer-screening

Force, U. S. P. S. T. (2021). Final Recommendation Statement Lung Cancer: Screening. Retrieved from https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/lung-cancer-screening

Foundation, N. K. Diabetes and Chronic Kideny Disease. Retrieved from https://www.kidney.org/news/newsroom/factsheets/Diabetes-And-CKD

Foundation, P. s. Statistics. Retrieved from https://www.parkinson.org/Understanding-Parkinsons/Statistics

Foundation, S. C. Skin Cancer Facts & Statistics. Retrieved from https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

Gerke, S., Minssen, T., & Cohen, G. (2020). Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare. In A. Bohr & K. Memarzadeh (Eds.), *Artificial Intelligence in Healthcare* (pp. 295-336): Academic Press.

Giavina-Bianchi, M., de Sousa, R. M., Paciello, V. Z. A., Vitor, W. G., Okita, A. L., Proa, R., . . . Machado, B. S. (2021). Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting. *PLoS One, 16*(9), e0257006. doi:10.1371/journal.pone.0257006

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr Psychiatry Rep, 21*(11), 116. doi:10.1007/s11920-019-1094-0

He, J., Cao, T., Xu, F., Wang, S., Tao, H., Wu, T., . . . Chen, J. (2020). Artificial intelligence-based screening for diabetic retinopathy at community hospital. *Eye (Lond), 34*(3), 572-576. doi:10.1038/s41433-019-0562-4

Health, N. I. o. What are some common complications of pregnancy? Retrieved from https://www.nichd.nih.gov/health/topics/pregnancy/conditioninfo/complications

Heath, S. (2016). How Does Health IT Interoperability Affect Patient Safety? Retrieved from https://ehrintelligence.com/news/how-does-health-it-interoperability-affect-patient-safety

Heo, J., Yoon, J. G., Park, H., Kim, Y. D., Nam, H. S., & Heo, J. H. (2019). Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. *Stroke, 50*(5), 1263-1265. doi:10.1161/STROKEAHA.118.024293

HIMSS. (2021). Interoperability in Healthcare. Retrieved from https://www.himss.org/resources/interoperability-healthcare#Part1

Hsiao, Y. J., Wen, Y. C., Lai, W. Y., Lin, Y. Y., Yang, Y. P., Chien, Y., . . . Jheng, Y. C. (2021). Application of artificial intelligence-driven endoscopic screening and diagnosis of gastric cancer. *World J Gastroenterol, 27*(22), 2979-2993. doi:10.3748/wjg.v27.i22.2979

IBM. Machine Learning. Retrieved from https://www.ibm.com/cloud/learn/machine-learning

IBM. Natural Language Processing (NLP). Retrieved from https://www.ibm.com/cloud/learn/machine-learning

Institute, T. N. E. Diabetic Retinopathy. Retrieved from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy

Jia, H., Li, L., Li, W., Hou, T., Ma, H., Yang, Y., . . . Chen, M. (2019). Impact of Healthcare-Associated Infections on Length of Stay: A Study in 68 Hospitals in China. *Biomed Res Int, 2019*, 2590563. doi:10.1155/2019/2590563

Joy Mathew, C., David, A. M., & Joy Mathew, C. M. (2020). Artificial Intelligence and its future potential in lung cancer screening. *EXCLI J, 19*, 1552-1562. doi:10.17179/excli2020-3095

Karami, K., Akbari, M., Moradi, M. T., Soleymani, B., & Fallahi, H. (2021). Survival prognostic factors in patients with acute myeloid leukemia using machine learning techniques. *PLoS One, 16*(7), e0254976. doi:10.1371/journal.pone.0254976

Kaushal, A., Altman, R., Langlotz C. (2020). Health Care AI Systems Are Biased.

. Retrieved from https://www.scientificamerican.com/article/health-care-ai-systems-are-biased/

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Med, 17*(1), 195. doi:10.1186/s12916-019-1426-2

Keskinbora, K. H. (2019). Medical ethics considerations on artificial intelligence. *J Clin Neurosci, 64*, 277-282. doi:10.1016/j.jocn.2019.03.001

Kickingereder, P., Isensee, F., Tursunova, I., Petersen, J., Neuberger, U., Bonekamp, D., . . . Maier-Hein, K. H. (2019). Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. *Lancet Oncol, 20*(5), 728-740. doi:10.1016/S1470-2045(19)30098-1

Kroenke, K., Wu, J., Yu, Z., Bair, M. J., Kean, J., Stump, T., & Monahan, P. O. (2016). Patient Health Questionnaire Anxiety and Depression Scale: Initial Validation in Three Clinical Trials. *Psychosom Med, 78*(6), 716-727. doi:10.1097/PSY.0000000000000322

Lai, Q., Spoletini, G., Mennini, G., Laureiro, Z. L., Tsilimigras, D. I., Pawlik, T. M., & Rossi, M. (2020). Prognostic role of artificial intelligence among patients with hepatocellular cancer: A systematic review. *World J Gastroenterol, 26*(42), 6679-6688. doi:10.3748/wjg.v26.i42.6679

Lee, D., & Yoon, S. N. (2021). Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges. *Int J Environ Res Public Health, 18*(1). doi:10.3390/ijerph18010271

Li, A. Y., & Elliot, N. (2019). Natural language processing to identify ureteric stones in radiology reports. *J Med Imaging Radiat Oncol, 63*(3), 307-310. doi:10.1111/1754-9485.12861

Luo, Y., Wang, Z., & Wang, C. (2021). Improvement of APACHE II score system for disease severity based on XGBoost algorithm. *BMC Med Inform Decis Mak, 21*(1), 237. doi:10.1186/s12911-021-01591-x

Makino, M., Yoshimoto, R., Ono, M., Itoko, T., Katsuki, T., Koseki, A., . . . Suzuki, A. (2019). Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. *Sci Rep*, *9*(1), 11862. doi:10.1038/s41598-019-48263-5

Management, H. D. Silos that limit patient record access are hurting care, researchers say. Retrieved from https://www.healthdatamanagement.com/articles/silos-that-limit-patient-record-access-are-hurting-care-researchers-say

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., . . . Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89-94. doi:10.1038/s41586-019-1799-6

McQuitty, O. (2020). The data elephant in the room. Retrieved from https://www.hospitaltimes.co.uk/the-data-elephant-in-the-room/

Medicine, J. H. Age-Related Macular Degeneration (AMD). Retrieved from https://www.hopkinsmedicine.org/health/conditions-and-diseases/agerelated-macular-degeneration-amd

Mello, M. M., Adler-Milstein, J., Ding, K. L., & Savage, L. (2018). Legal Barriers to the Growth of Health Information Exchange-Boulders or Pebbles? *Milbank Q, 96*(1), 110-143. doi:10.1111/1468-0009.12313

Montazeri, M., Montazeri, M., Montazeri, M., & Beigzadeh, A. (2016). Machine learning models in breast cancer survival prediction. *Technol Health Care, 24*(1), 31-42. doi:10.3233/THC-151071

Narkhede, S. (2018). Understanding AUC - ROC Curve. Retrieved from https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

Oliveira, C. R., Niccolai, P., Ortiz, A. M., Sheth, S. S., Shapiro, E. D., Niccolai, L. M., & Brandt, C. A. (2020). Natural Language Processing for Surveillance of Cervical and Anal Cancer and Precancer: Algorithm Development and Split-Validation Study. *JMIR Med Inform, 8*(11), e20826. doi:10.2196/20826

Opthalamology, A. A. o. What Is Macular Degeneration? Retrieved from https://www.aao.org/eye-health/diseases/amd-macular-degeneration

Organization, W. H. Maternal mortality: Levels and trends 2000 to 2017. Retrieved from https://www.who.int/reproductivehealth/publications/maternal-mortality-2000-2017/en/

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *J Glob Health, 9*(2), 010318. doi:10.7189/jogh.09.020318

Pancreatic Cancer: Statistics. Retrieved from https://www.cancer.net/cancer-types/pancreatic-cancer/statistics

Parthasarathy, G., Lopez, R., McMichael, J., & Burke, C. A. (2020). A natural language-based tool for diagnosis of serrated polyposis syndrome. *Gastrointest Endosc, 92*(4), 886-890. doi:10.1016/j.gie.2020.04.077

Prevention, A. F. f. S. Suicide Statistics. Retrieved from https://afsp.org/suicide-statistics/

Prevention, C. f. D. C. a. Stroke Facts. Retrieved from https://www.cdc.gov/stroke/facts.htm

Prevention, C. f. D. C. a. What are the Risk Factors for Breast Cancer? Retrieved from https://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm

Prevention, C. f. D. C. a. What is a Mammogram? Retrieved from https://www.cdc.gov/cancer/breast/basic_info/mammograms.htm

Prevention, C. f. D. C. a. (2014). Types of Healthcare-associated Infections. Retrieved from https://www.cdc.gov/hai/infectiontypes.html

Prevention, C. f. D. C. a. (2021). Aortic Aneurysm. Retrieved from https://www.cdc.gov/heartdisease/aortic_aneurysm.htm

Reisman, M. (2017). EHRs: The Challenge of Making Electronic Data Usable and Interoperable. *P T, 42*(9), 572-575. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28890644

Report, B. Risks and remedies for artificial intelligence in health care. Retrieved from https://www.brookings.edu/research/risks-and-remedies-for-artificial-intelligence-in-health-care/

Review, H. B. The Risks of Using AI to Interpret Human Emotions. Retrieved from https://hbr.org/2019/11/the-risks-of-using-ai-to-interpret-human-emotions

Safari, S., Baratloo, A., Elfil, M., & Negida, A. (2015). Evidence Based Emergency Medicine Part 2: Positive and negative predictive values of diagnostic tests. *Emerg (Tehran), 3*(3), 87-88. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26495390

Schaffter, T., Buist, D. S. M., Lee, C. I., Nikulin, Y., Ribli, D., Guan, Y., . . . Jung, H. (2020). Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open, 3*(3), e200265. doi:10.1001/jamanetworkopen.2020.0265

Shahi, N., Phillips, R., Meier, M., Smith, D., Leopold, D., Recicar, J., & Moulton, S. (2020). The true cost of child abuse at a level 1 pediatric trauma center. *J Pediatr Surg*, *55*(2), 335-340. doi:10.1016/j.jpedsurg.2019.10.024

Shahi, N., Shahi, A. K., Phillips, R., Shirek, G., Lindberg, D. M., & Moulton, S. L. (2021). Using deep learning and natural language processing models to detect child physical abuse. *J Pediatr Surg, 56*(12), 2326-2332. doi:10.1016/j.jpedsurg.2021.03.007

Society, A. C. Cervical Cancer. Retrieved from https://www.cancer.org/cancer/cervical-cancer.html

Society, A. C. Gatric Cancer. Retrieved from https://www.cancer.org/cancer/stomach-cancer.html

Society, A. C. Lung Cancer Risk Factors. Retrieved from https://www.cancer.org/cancer/lung-cancer/causes-risks-prevention/risk-factors.html

Society, A. C. Nasopharyngeal Cancer. Retrieved from https://www.cancer.org/cancer/nasopharyngeal-cancer.html

Society, A. C. Ovarian Cancer Risk Factors. Retrieved from https://www.cancer.org/cancer/ovarian-cancer/causes-risks-prevention/risk-factors.html

Society, A. C. Risk Factors for Acute Myeloid Leukemia (AML). Retrieved from https://www.cancer.org/cancer/acute-myeloid-leukemia/causes-risks-prevention/risk-factors.html

Society, A. C. Signs and Symptoms of Lung Cancer. Retrieved from https://www.cancer.org/cancer/lung-cancer/detection-diagnosis-staging/signs-symptoms.html

Society, A. C. Treatment of Invasive Epithelial Ovarian Cancers, by Stage. Retrieved from https://www.cancer.org/cancer/ovarian-cancer/treating/by-stage.html

Society, A. C. Treatment Options for Cervical Cancer, by Stage

. Retrieved from https://www.cancer.org/cancer/cervical-cancer/treating/by-stage.html

Society, A. C. (2020). Risk Factors for Anal Cancer. Retrieved from https://www.cancer.org/cancer/anal-cancer/causes-risks-prevention/risk-factors.html

Society, A. C. (2021a). How Common Is Breast Cancer? Retrieved from https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html

Society, A. C. (2021b). Key Statistics About Liver Cancer. Retrieved from https://www.cancer.org/cancer/liver-cancer/about/what-is-key-statistics.html

Society, A. C. (2021c). Key Statistics for Lung Cancer. Retrieved from https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html

Society, A. C. (2021d). Key Statistics for Ovarian Cancer. Retrieved from https://www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html

Society, A. C. (2022). Stomach Cancer Survival Rates. Retrieved from https://www.cancer.org/cancer/stomach-cancer/detection-diagnosis-staging/survival-rates.html

Solutions, A. Homage to John McCarthy, the Father of Artificial Intelligence (AI). Retrieved from https://www.artificial-solutions.com/blog/homage-to-john-mccarthy-the-father-of-artificial-intelligence

Soni, A., Amrhein, B., Baucum, M., Paek, E. J., & Khojandi, A. (2021). Using Verb Fluency, Natural Language Processing, and Machine Learning to Detect Alzheimer's Disease. *Annu Int Conf IEEE Eng Med Biol Soc, 2021*, 2282-2285. doi:10.1109/EMBC46164.2021.9630371

Sulmasy, L. S., Lopez, A. M., Horwitch, C. A., Professionalism, A. C. o. P. E., & Human Rights, C. (2017). Ethical Implications of the Electronic Health Record: In the Service of the Patient. *J Gen Intern Med, 32*(8), 935-939. doi:10.1007/s11606-017-4030-1

Times, P. Who Owns Patients' Data? Retrieved from https://www.pharmacytimes.com/view/who-owns-patients-data

Tvardik, N., Kergourlay, I., Bittar, A., Segond, F., Darmoni, S., & Metzger, M. H. (2018). Accuracy of using natural language processing methods for identifying healthcare-associated infections. *Int J Med Inform, 117*, 96-102. doi:10.1016/j.ijmedinf.2018.06.002

Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., . . . Stroke Statistics, S. (2021). Heart Disease and Stroke Statistics-2021 Update: A Report From the American Heart Association. *Circulation, 143*(8), e254-e743. doi:10.1161/CIR.0000000000000950

Wang, S. V., Rogers, J. R., Jin, Y., Bates, D. W., & Fischer, M. A. (2017). Use of electronic healthcare records to identify complex patients with atrial fibrillation for targeted intervention. *J Am Med Inform Assoc, 24*(2), 339-344. doi:10.1093/jamia/ocw082

Wang, Y., Liu, S., Wang, Z., Fan, Y., Huang, J., Huang, L., . . . Zhou, F. (2021). A Machine Learning-Based Investigation of Gender-Specific Prognosis of Lung Cancers. *Medicina (Kaunas), 57*(2). doi:10.3390/medicina57020099

Wu, C. C., Hsu, W. D., Islam, M. M., Poly, T. N., Yang, H. C., Nguyen, P. A., . . . Li, Y. J. (2019). An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. *Comput Methods Programs Biomed*, *173*, 109-117. doi:10.1016/j.cmpb.2019.01.013