

Review

Machine-Learning-Based Prediction Modelling in Primary Care: State-of-the-Art Review

Adham H. El-Sherbini ¹, Hafeez Ul Hassan Virk ², Zhen Wang ^{3,4}, Benjamin S. Glicksberg ⁵
and Chayakrit Krittanawong ^{6,*}

¹ Faculty of Health Sciences, Queen's University, Kingston, ON K7L 3N6, Canada

² Harrington Heart & Vascular Institute, Case Western Reserve University, University Hospitals Cleveland Medical Center, Cleveland, OH 44115, USA

³ Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN 55901, USA

⁴ Division of Health Care Policy and Research, Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55901, USA

⁵ The Hasso Plattner Institute for Digital Health at the Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁶ Cardiology Division, NYU School of Medicine and NYU Langone Health, New York, NY 10016, USA

* Correspondence: chayakrit.krittanawong@nyulangone.org

Abstract: Primary care has the potential to be transformed by artificial intelligence (AI) and, in particular, machine learning (ML). This review summarizes the potential of ML and its subsets in influencing two domains of primary care: pre-operative care and screening. ML can be utilized in preoperative treatment to forecast postoperative results and assist physicians in selecting surgical interventions. Clinicians can modify their strategy to reduce risk and enhance outcomes using ML algorithms to examine patient data and discover factors that increase the risk of worsened health outcomes. ML can also enhance the precision and effectiveness of screening tests. Healthcare professionals can identify diseases at an early and curable stage by using ML models to examine medical pictures, diagnostic modalities, and spot patterns that may suggest disease or anomalies. Before the onset of symptoms, ML can be used to identify people at an increased risk of developing specific disorders or diseases. ML algorithms can assess patient data such as medical history, genetics, and lifestyle factors to identify those at higher risk. This enables targeted interventions such as lifestyle adjustments or early screening. In general, using ML in primary care offers the potential to enhance patient outcomes, reduce healthcare costs, and boost productivity.

Keywords: artificial intelligence; machine learning; deep learning; primary care



Citation: El-Sherbini, A.H.; Hassan Virk, H.U.; Wang, Z.; Glicksberg, B.S.; Krittanawong, C.

Machine-Learning-Based Prediction Modelling in Primary Care:

State-of-the-Art Review. *AI* **2023**, *4*, 437–460. <https://doi.org/10.3390/ai4020024>

Academic Editors: Tim Hulsen and Demos T. Tsahalidis

Received: 14 March 2023

Revised: 8 May 2023

Accepted: 16 May 2023

Published: 23 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence (AI) is a field of study that attempts to replicate natural human intelligence in machines [1]. The machines can then independently perform activities that would otherwise require human intelligence. AI can be broken down into several subsets, such as machine learning (ML) and deep learning (DL) [2]. ML makes a software application more accurate in predicting outcomes by feeding it with data rather than explicit programming. Comparatively, DL, a subset of ML, builds a hierarchy of knowledge based on learning from examples. These fundamental ideas of AI are utilized to develop analytic models to turn this productive technology into practice. Since its introduction in the 1950s, AI has made significant strides in manufacturing; sports analytics; autonomous vehicle; and more recently, primary care and preventive medicine [3].

Primary care and preventive medicine, otherwise expressed as day-to-day healthcare practices including outpatient settings, are a growing sector in the realms of AI and computer science. Although AI has endless applications in healthcare, particular sectors of primary care have been more progressive and accepting of AI and its potential. For

instance, the Forward clinic is a primary care service incorporating standard doctor-led programs with technology to provide a more inclusive and long-term care [3]. The addition of the technology allows for 24/7 monitoring, skin cancer screening, testing of genes, and biometric monitoring. As with all AI interventions, the Forward clinic endures multiple challenges, such as additional physician training and fees. Although the Forward clinic is just a singular example of how AI can be integrated into primary care, AI's implementation into primary care can be further broken down into sections of healthcare, such as pre-operative care and screening. This review summarizes AI's, specifically ML's, short yet productive impact on primary care and preventive medicine and aims to inform primary care physicians about the potential integration of ML (Figure 1 and Table 1).

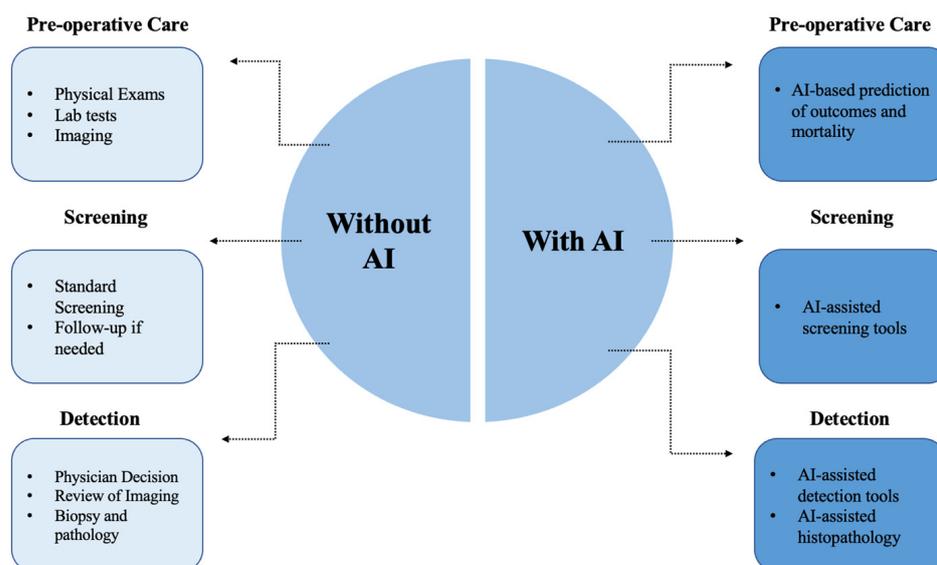


Figure 1. Current methods vs. AI-assisted methods in primary care. Figure Description: AI has the potential to assist current primary care methods in three domains: pre-operative care, screening, and detection. In pre-operative care, this includes using AI for predictions of outcomes and mortality. For screening, AI serves a prominent role in screening tools for numerous diseases. Similarly, AI can be used for real-time detection tools and AI-assisted histopathology tools.

Table 1. Abbreviations.

Name	Abbreviation
Acute kidney injury	AKI
Adaptive boosting	ADA
Age-related macular degeneration	AMD
Artificial intelligence	AI
Atherosclerotic cardiovascular disease	ACSVD
Atrial fibrillation	AF
Blood pressure	BP
Chronic kidney disease	CKD
Chronic obstructive pulmonary disease	COPD
Convolutional neural network	CNN
Coronary artery calcium score	CACS
Coronary artery disease	CAD
Decision tree	DT
Deep learning	DL
Deep neural network	DNN
Deep vein thrombus	DVT
Diabetes mellitus	DM
Electronic health records	EHR
Extreme gradient boosting	XGB

Table 1. *Cont.*

Name	Abbreviation
Familial hypercholesterolemia	FH
Generative adversarial network	GAN
Gradient boosting	GB
Gradient boosting tree	GBT
Heart failure	HF
Human immunodeficiency virus	HIV
K-nearest neighbors	KNN
Logistic regression	LR
Low-density lipoprotein	LDL
Machine learning	ML
Neural network	NN
Obstructive sleep apnea syndrome	OSAS
Photoplethysmogram	PPG
Potential pre-exposure prophylaxis	PrEP
Pulmonary embolism	PE
Pulmonary hypertension	PH
Random forest	RF
Support vector machine	SVM
Urinary tract infection	UTI

2. Pre-Operative Care

Pre-operative risk prediction and management have been promising areas of AI research and its application. PubMed and Google Scholar were searched using keywords for English literature published from inception to December 2022 (Figure 2). Studies were included if they reported outcomes regarding the effectiveness of ML models in pre-operative care or similar domains. Studies have utilized AI to predict mortality and postoperative complications. Such applications are necessary for clinical decision-making, forethought of healthcare resources such as ICU beds, the cost of the patient, and the possible need for transition of care [4]. Typically, researchers utilize a designated number of electronic health records (EHR) to train the analytic model and the remainder to test it. For instance, Chiew et al. utilized EHRs to predict post-surgical mortality in a tertiary academic hospital in Singapore [5]. The study compared five candidate models (Random Forest (RF), Adaptive Boosting (ADA), Gradient Boosting (GB), and Support Vector Machine (SVM)) and found that all GB was the greatest performing model (specificity (0.98), sensitivity (0.50), PPV (0.20), F1 score (0.28), and AUROC (0.96)). Five other studies by Fernandes et al., Jalai et al., COVIDSurg Collaborative, Sahara et al., and Pfitzner et al. have also evaluated how differing types of analytic models (Logistic Regression (LR), RF, Neural Network (NN), SVM, Extreme GB (XGB), Decision Tree (DT), GB, Deep Neural Network (DNN), GRU, and classification tree) can predict postoperative mortality [6–10]. The patient population included those undergoing cardiac surgery, pancreatic surgery, or hepatopancreatic surgery or those infected with SARS-CoV-2. Of the studies undergoing cardiac surgery, the selected ML models were good predictors of mortality and prolonged length of stay. In Fernandes et al., when utilizing pre-operative and intra-operative risk factors alongside intraoperative hypotension, XGB was the best performing model (AUROC (0.87), PPV (0.10), specificity (0.85), and sensitivity (0.71) [6]. In the other study by Jalai et al., deep neural network (DNN) was the best performing of the five models (accuracy (89%), F-score (0.89), and AUROC (0.95)) [7]. Neither study compared its models with established pre-operative risk scores, such as the Revised Cardiac Risk Index or Gupta score. Similarly, Pfitzner et al. used pre-, intra-, and short-term post-operative data on a number of models to assess its ability to predict pre-operative risk for those undergoing pancreatic surgery [8]. The study found maximum AUPRCs of 0.53 for postoperative complications and 0.51 for postoperative mortality, with LR as the best model. As for those undergoing hepatopancreatic surgery, Sahara et al. found that the classification tree model better

predicted 30-day unpredicted deaths than the traditional American College of Surgeons National Surgery Quality Improvement Program surgical risk calculator [9]. Finally, a COVIDSurg Collaborative study that generated 78 AI models found that when combining an LR model with four features (ASA grade, RCRI, age, and pre-op respiratory support), an AUC of 0.80 in the testing dataset was achieved. This generated model was the best performing in predicting postoperative mortality among those infected with SARS-CoV-2 [10]. Ultimately, ML models present great promise in its integration into pre-operative care, particularly for simplifying pre-operative evaluations, as observed in Figure 3.

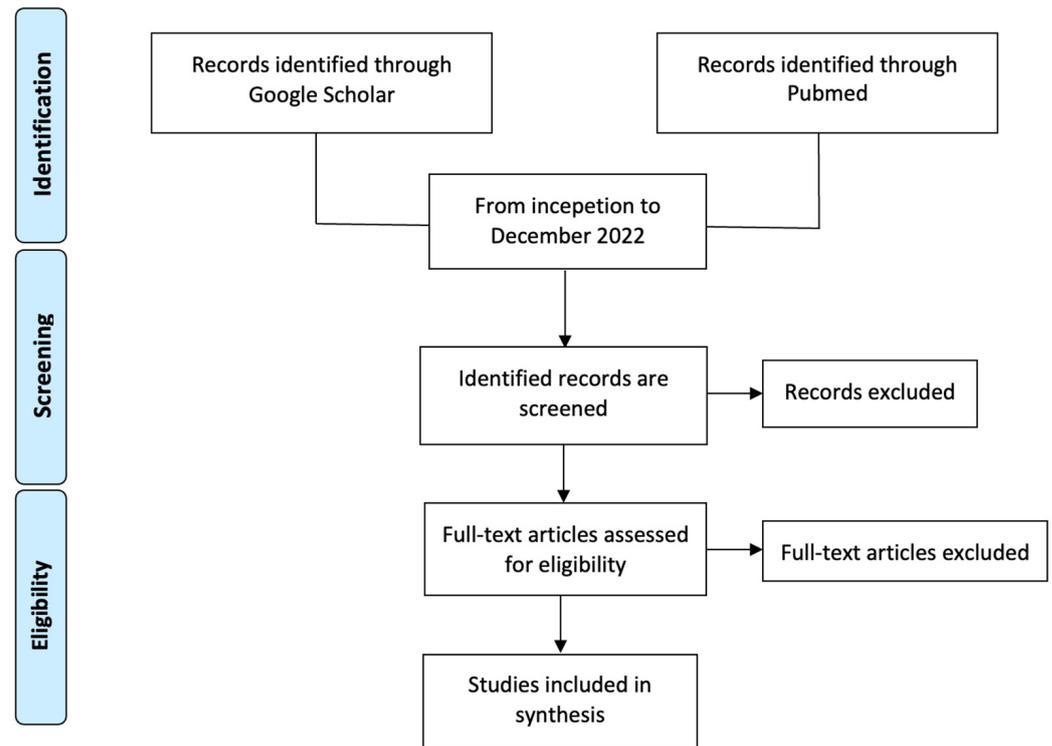


Figure 2. Literature Search Method. Figure Description: PubMed and Google Scholar were searched using keywords for English literature published from inception to December 2022. Observational studies, case-control studies, cohort studies, clinical trials, meta-analyses, reviews, and guidelines were reviewed.

Post-Operative Complications

Other pre-operative risk prediction objectives include assessing models on postoperative complications [11–13]. These studies have evaluated how varying ML models (SVM, LR, RF, GBT, DNN, GBT, and XGB) can predict a number of post-operative complications. One study utilized electronic anesthesia records (pre-operative and intra-operative data) to predict deep vein thrombus (DVT), delirium, pulmonary embolism, acute kidney injury (AKI), and pneumonia [11]. GBT was the most promising model, with AUROC scores of 0.905 (pneumonia), 0.848 (AKI), 0.881 (DVT), 0.831 (pulmonary embolism), and 0.762 (delirium). Similarly, Corey et al. utilized EHR data, including 194 clinical features, to train ML models on 14 postoperative complications [12]. Amongst the models, AUC scores ranged from 0.747 to 0.924, with the Lasso penalized regression being the best performing (sensitivity (0.775), specificity (0.749), and PPV (0.362)). Comparably, Bonde et al. trained three multi-label DNNs to compete against traditional surgical risk prediction systems on post-operative complications [13]. The mean AUCs for the test dataset on models 1, 2, and 3 were 0.858, 0.863, and 0.874, all of which outperformed the ACS-SRC predictors. Ultimately, ML methods appear to be high-performing for predicting post-operative complications, but additional studies comparing models are required to validate the findings.

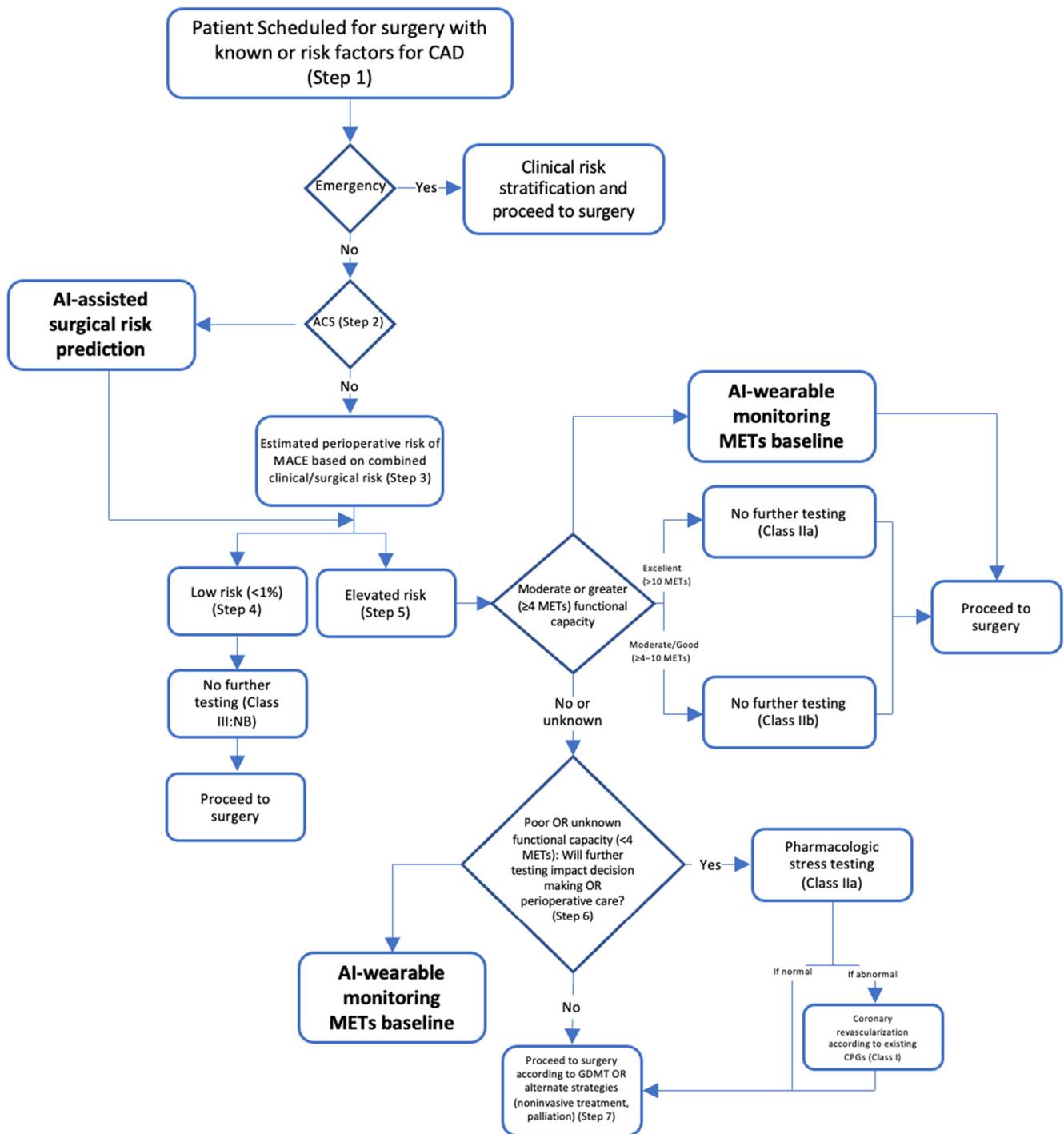


Figure 3. Example of AI in pre-operative evaluation. Figure Description: The integration of AI into pre-operative care allows for the refinement of more effective guidelines. For instance, in Figure 2, current guidelines recommend a seven-step pre-operative evaluation before surgery for patients with CAD risk factors. In this process, AI could be utilized to provide risk prediction and MET monitoring through wearable technology, ultimately cultivating a more straightforward process.

3. Screening

The applications of AI in screening are by far the most precedented. PubMed and Google Scholar were searched from inception to December 2022, and the databases were searched for studies investigating the role of ML in screening for several diseases and disorders.

3.1. Hypertension

One of these leading domains is hypertension, where studies have assessed the risk of hypertension and predicted resistant hypertension while concurrently estimating blood pressure (BP). Zhao et al. compared four analytical models (RF, CatBoost, MLP neural network, and LR) in predicting the risk of hypertension based on data from physical examinations [14]. RF was the best-performing model with an AUC of 0.92, an accuracy of 0.82, a sensitivity of 0.83, and a specificity of 0.81. In addition, no clinical or genetic data was utilized for training the models. Similarly, Alkaabi et al. utilized ML models (DT, RF, and LR) to assess the risk of developing hypertension in a more effective manner [15]. RF was the best-performing model (accuracy (82.1%), PPV (81.4%), sensitivity (82.1%), and AUC (86.9)). Clinical factors, such as education level, tobacco use, abdominal obesity, age, gender, history of diabetes, consumption of fruits and vegetables, employment, physical activity, mother's history of high BP, and history of high cholesterol, were all significant predictors of hypertension. Ye et al. investigated an XGBoost model that had AUC scores of 0.917 (retrospective) and 0.870 (prospective) in predicting hypertension. Similarly, LaFreniere et al. investigated an NN model which had 82% accuracy in predicting hypertension given the chosen risk factors [16,17]. Regarding BP, Khalid et al. compared three ML models (regression tree, SVM, and MLR) in estimating BPs from pulse waveforms derived from photoplethysmogram (PPG) signals [18]. The regression tree achieved the best systolic and diastolic BP accuracy, -0.1 ± 6.5 mmHg and -0.6 ± 5.2 mmHg, respectively. In summary, ML appears to be an effective tool for predicting hypertension and BP, though its clinical utility remains to be delineated, since hypertension can be diagnosed through non-invasive procedures.

3.2. Hypercholesterolemia

AI applications on hypercholesterolemia have outputted similar findings, as seen in Myers et al. [19]. Using data on diagnostic and procedure codes, prescriptions, and laboratory findings, the FIND FH model was trained on large healthcare databases to diagnose familial hypercholesterolemia (FH). The model achieved a PPV of 0.85, a sensitivity of 0.45, an AURPC of 0.55, and an AUROC score of 0.89. This model effectively identified those with FH for individuals at high risk of early heart attack and stroke. Comparatively, Pina et al. evaluated the accuracy of three ML models (CT, GBM, and NN) when trained on genetic tests to detect FH-causative genetic mutations [20]. All three models outperformed the clinical standard Dutch Lipid score in both cohorts. Similar findings have been produced for hyperlipidemia, where Liu et al. trained an LSTM network on 500 EHR samples [21]. The model achieved an ACC score of 0.94, an AUC score of 0.974, a sensitivity of 0.96, and a specificity of 0.92. Regarding low-density lipoproteins (LDLs), Tsigalou et al. and Cubukcu et al. concluded that ML models were productive alternatives to direct determination and equations [22,23]. In both studies, ML models (MLR, DNN, ANN, LR, and GB trees) outperformed the traditional equations: the Friedewald and Martin-Hopkins formulas. Although the researched algorithms show great potential, additional studies are warranted to validate these conclusions.

3.3. Cardiovascular Disease

Arguably, the largest field of primary care in which AI has made significant strides is predicting and assessing cardiovascular risk. As cardiovascular diseases are the leading cause of death globally, any advancements in risk prediction and early diagnosis are of substance. In 2017, Weng et al. compared four ML models (RF, LR, GB, and NN) in predicting cardiovascular risk through EHR [24]. The AUC scores of RF, LR, GB, and NN were 0.745, 0.760, 0.761, and 0.764, respectively. The study concluded that the applications of ML in cardiovascular risk prediction significantly improved the accuracy. Zhao et al. reproduced a similar study with LR, RF, GB, CNN, and LSTM trained on longitudinal EHR and genetic data [25]. The event prediction was far better using the longitudinal feature for a 10-year CVD prediction. Kusunose et al. applied a CNN to identify those at risk of heart failure (HF) from a cohort of pulmonary hypertension

(PH) patients using chest x-rays [26]. The AUC scores of AI, chest x-rays, and human observers were 0.71, 0.60, and 0.63, respectively. In a unique perspective, Moradi et al. employed generative adversarial networks (GANs) for data augmentation on chest x-rays to assess its accuracy in detecting cardiovascular abnormalities when a CNN model was trained on it [27]. The GAN data augmentation outperformed traditional and no data augmentation scenarios on normal and abnormal chest X-ray images with accuracies of 0.8419, 0.8312, and 0.8193, respectively. Studies have also compared ML models relative to traditional risk scores, such as a study by Ambale-Venkatesh et al. [28]. A random survival forest model was assessed in its prediction of six cardiovascular outcomes compared with the Concordance index and Brier score. The model outperformed traditional risk scores (decreased Brier score by 10–25%), and age was the most significant predictor. Similarly, Alaa et al. compared an AutoPrognosis ML model with an established risk score (Framingham score), a Cox PH model with conventional risk factors, and a Cox PH model with all 473 variables (UK Biobank) [29]. The AUROC scores were 0.774, 0.724, 0.734, and 0.758, respectively. Pfohl et al. developed a “fair” atherosclerotic cardiovascular disease (ACSVD) risk prediction tool through EHR data [30]. The experiment ran through four experiments (standard, EQ_{race} , EQ_{gender} , and EQ_{age}) and achieved AUROC scores of 0.773, 0.742, 0.743, and 0.694, respectively. The tool has reduced discrepancies across races, genders, and ages in the prediction of ACSVD. Generally, AI can aid in mitigating gaps in ACSVD risk prevention guidelines, as observed in Figure 4. In the United States alone, one in every three patients undergoing elective cardiac catheterization is diagnosed with obstructive coronary artery disease (CAD). This begs the question of new methodologies to better diagnose the population. Al’Aref et al. assessed how applying an XGBoost model on Coronary Computed Tomography Angiography can predict obstructive CAD using clinical factors [31]. The ML model achieved an AUC score of 0.773, but more notably, when combined with the coronary artery calcium score (CACs), the AUC score was 0.881. Therefore, an ML model and CACS may accurately predict the presence of obstructive CAD. Based on the present literature, AI models screen effectively and predict cardiovascular risks while predominantly outperforming established risk scores.

3.4. Eye Disorders and Diseases

Another area of primary care that has used ML is vision-centric diseases, such as diabetic retinopathy, glaucoma, and age-related macular degeneration (AMD). Ting et al. assessed AI’s metrics in this sector by training a DL system on retinal images (76,370 images of diabetic retinopathy, 125,189 images of possible glaucoma, and 72,610 images of AMD) [32]. For referable diabetic retinopathy, the model achieved an AUC of 0.936, a sensitivity of 0.905, and a specificity of 0.916. As for vision-threatening retinopathy, the AUC was 0.958, sensitivity was 1.00, and specificity was 0.911. For possible glaucoma images, the model achieved an AUC of 0.942, a sensitivity of 0.964, and a specificity of 0.872. Finally, the model on AMD testing retinal images achieved an AUC of 0.931, a sensitivity of 0.923, and a specificity of 0.887. Retinal fundus images can also be used by AI models to extract further information, such as predicting cardiovascular risk factors in the case of the study by Poplin et al. [33]. After training the model on 284,445 and validating it on two datasets, the model could predict age (mean absolute error (MAE) within 3.26 years), gender (AUC 0.97), smoking status (AUC 0.71), systolic blood pressure (MAE within 11.23 mmHG), and major adverse cardiac events (AUC 0.70). In another study, Kim et al. utilized retinal fundus images for training a CNN model to predict age and sex [34]. The MAE for patients, those with hypertension, those with diabetes mellitus (DM), and smokers were 3.06 years, 3.46 years, 3.55 years, and 2.65 years, respectively. Ultimately, well-trained ML models appear to be effective in predicting eye diseases.

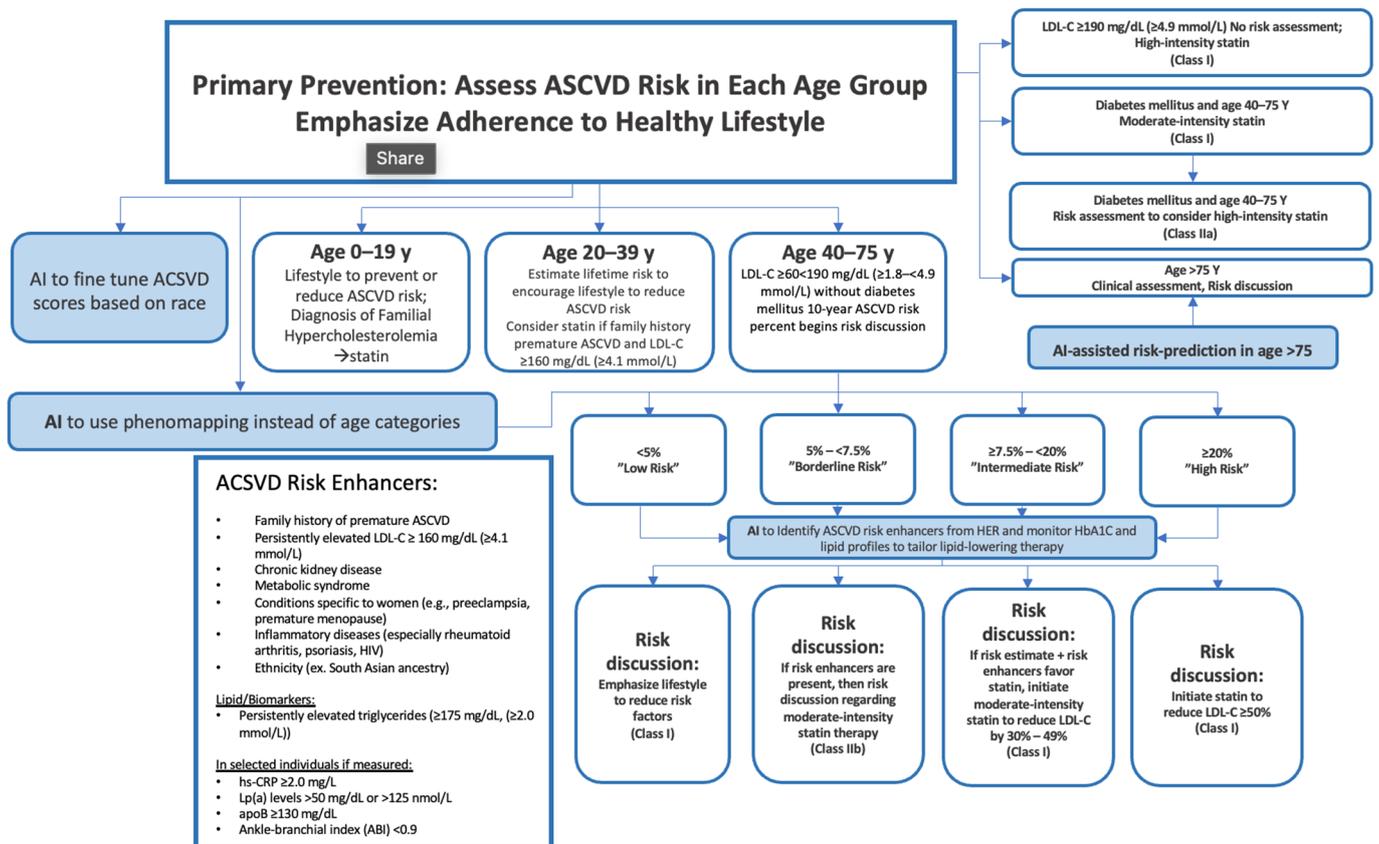


Figure 4. Example of AI in ASCVD assessment. Figure Description: ASCVD risk assessment is exceptionally extensive and varies significantly based on age groups. Although the guidelines are thorough, AI has the potential to address potential gaps in the evaluation. For instance, AI can provide risk prediction for individuals > 75 and it could fine-tune ACSVD scores based on race. AI could also detect risk enhancers of ASCVD based on HbA1C monitoring, EHR, and lipid profiles. This may allow for appropriate adjustments to lipid-lowering therapy. AI also has the potential to use phenomapping instead of age categories, allowing for stronger classification.

3.5. Diabetes

More than 400 million individuals globally are diagnosed with DM. AI's implementation into primary care has been shown to be effective when targeting this widespread disease's risk prediction. In one study, Alghamdi et al. used medical records of cardiorespiratory fitness to train and compare five models (DT, naïve bayes, LR, logistic model tree, and RF) in predicting DM. When RF, logistic model tree, and naïve bayes were ensembled with the developed predictive model classifier, a maximum AUC (0.92) was achieved. Similarly, through administrative data, Ravaut et al. trained a GB decision tree on 1,657,395 patients to predict T2DM 5 years prior to onset [35]. While validating the model on 243,442 patients and testing it on 236,506 patients, an AUC score of 0.8026 was achieved. In another study, Ravaut et al. also assessed if a GB decision tree can predict adverse complications of diabetes, including retinopathy, tissue infection, hyper/hypoglycemia, amputation, and cardiovascular events [36]. After being trained (1,029,366 patients), validated (272,864 patients), and tested (265,406 patients) on administrative data, the model achieved an AUC score of 0.777. To support the conclusion on DM, Deberneh et al. found reasonably good accuracies in a Korean population, with DT (77.87%), LR (76.13%), and ANN having the lowest accuracy (73.23%) [37]. In Alhassan et al., when predicting T2DM, the LSTM and gated-recurrent unit outperformed MLP models with a 97.3% accuracy [38]. In India, Boutillier et al. attempted to find the best ML algorithm for predicting DM and hypertension in limited resource settings [39]. RF models had a higher prediction accuracy

than established UK and US scores, with an improved AUC score from 0.671 to 0.910 for diabetes and from 0.698 to 0.792 for hypertension. With the current evidence, ML methods appear to be exceptionally effective in predicting diabetes; however, there lacks discussion on the benefits of using ML over a simple blood draw.

3.6. Cancer

In 2020, cancer was responsible for nearly 10 million deaths globally, making it a hotspot for ML implementations and strategies in primary care [40]. Fortunately, ML models have been proven to have potential in the early diagnosis and screening of lung, cervical, colorectal, breast, and prostate cancer [41]. Regarding lung cancer, Ardilla et al. trained a DL algorithm on CT images to predict the risk of lung cancer in 6716 national trial cases [42]. The model achieved an AUC score of 0.944. Similarly, Gould et al. compared an ML model in predicting a future lung cancer diagnosis with the 2012 Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial risk model (mPLCOm2012) [43]. The novel algorithm outperformed the mPLCOm2012 in AUC scores (0.86 vs. 0.79) and sensitivity (0.401 vs. 0.279). Using NNs, Yeh et al. developed a model to screen patients at risk of lung cancer on EHR data [44]. For the overall population, the algorithm achieved an AUC score of 0.90 and 0.87 for patients over the age of 55 years. Guo et al. trained ML models on low-dose CT and found an accuracy of 0.6778, a F1 score of 0.6575, a sensitivity of 0.6252, and a specificity of 0.7357 [45]. More notably, the interactive pathways were BMI, DM, first smoke age, average drinks per month, years of smoking, year(s) since quitting smoking, sex, last dental visit, general health, insurance, education, last PAP test, and last sigmoidoscopy or colonoscopy. Concerning cervical cancer, CervDetect, a number of ML models that evaluate the risk of cervical cancer elements forming, has been a leader in this subject. In 2021, Mehmood et al. used cervical images to evaluate CervDetect and found a false negative rate of 100%, a false positive rate of 6.4%, an MSE error of 0.07111, and an accuracy of 0.936 [46]. Similarly, DeepCervix is another DL model that attempts to combat the high false-positive results in pap smear tests due to human error. Rahaman et al. trained DeepCervix, a hybrid deep fusion feature technique, on pap smear tests [47]. The DL-based model achieved accuracies of 0.9985, 0.9938, and 0.9914 for 2-class, 3-class, and 5-class classifications, respectively. Considering that 90% of cervical cancer is found in low-middle income settings, Bae et al. set out to implement an ML model onto endoscopic visual inspection following an application of acetic acid images [48]. Although resource-limited, the KNN model was the best performing, with an accuracy of 0.783, an AUC of 0.807, a specificity of 0.803, and a sensitivity of 0.75. In parallel, Wentzensen et al. developed a DL classifier with a cloud-based whole-slide imaging platform and trained it on P16/Ki-67 dual-stained (DS) slides for cervical cancer screening [48]. The model achieved a better specificity and equal sensitivity to manual DS and pap, resulting in lower positivity than manual DS and cytology. With respect to breast cancer screening, multiple studies have been conducted to achieve better accuracy in its diagnosis. Using screening mammograms, Shen et al. trained a DL algorithm on 1903 images and achieved an AUC of 0.88, and the four-model averaging improved the AUC score to 0.91 [49]. Similarly, using digital breast tomosynthesis images, Buda et al. achieved a sensitivity of 65% with a DL model for breast cancer screening [50]. Similarly, Haji Maghsoudi et al. developed Deep-LIBRA, an AI model trained on 15661 digital mammograms to estimate breast density and achieved an AUC of 0.612 [51]. The model had a strong agreement with the current gold standard. Another study by Ming et al. compared three ML models (MCMC GLMM, ADA, and RF) to the established BOADICEA model by training them on biennial mammograms [52]. When screening for lifetime risk of breast cancer, all models ($0.843 \leq \text{AUROC} \leq 0.889$) outperformed BOADICEA (AUROC = 0.639). Similar findings have been concluded in prostate cancer, where three studies (Perera et al., Chiu et al., and Bienecke et al.) compared numerous ML models (DNN, XGBoost, LightGBM, CatBoost, SVM, LR, RF, and multiplayer perceptron) [53–55]. Although all studies trained their respective models differently (PSA levels, prostate biopsy, or EHRs), all concluded that the ML algorithms were efficacious in

predicting prostate cancer. Ultimately, there appears to be a substantial body of literature supporting the effectiveness of ML methods in predicting different types of cancer.

3.7. Human Immunodeficiency Virus and Sexually Transmitted Diseases

Another sector of primary care requiring additional applications to assist in its diagnosis and screening is the human immunodeficiency virus (HIV) and sexually transmitted diseases (STDs). In 2021, Turbe et al. trained a DL model on the rapid diagnostic test to classify rapid HIV in rural South Africa [56]. Relative to traditional reports of accuracy varying between 80 and 97%, this model achieved an accuracy of 98.9%. Similarly, Bao et al. compared 5 mL models predicting HIV and STIs [57]. GBM was the best performing, with AUROC scores of 0.763, 0.858, 0.755, and 0.68 for HIV, syphilis, gonorrhea, and chlamydia, respectively. Another study, Marcus et al., developed and assessed an HIV prediction model to find potential pre-exposure prophylaxis (PrEP) patients [58]. Using EHR data to train the model, the study reported an AUC score of 0.84. In terms of future predictions, Elder et al. compared 6 mL algorithms when determining patients at risk of additional STIs within the next 24–48 months through previous EHR data [59]. The Bayesian Additive RT was the best-performing model with an AUROC score of 0.75 and a sensitivity of 0.915. A number of studies have also reported plausible applications of AI on urinary tract infections (UTIs). Gadalla et al. have assessed how AI models can identify predictors for a UTI diagnosis through training on potential biomarkers and clinical data from urine [60]. The study concluded that clinical information was the strongest predictor, with an AUC score of 0.72, a PPV of 0.65, an NPV of 0.79, and an F1 score of 0.69. Comparatively, in Taylor et al., vitals, lab results, medication, chief complaints, physical exam findings, and demographics were all utilized for training, validating, and testing a number of ML algorithms to predict UTIs in ED patients [61]. The AUC scores ranged from 0.826 to 0.904, with XGBoost being the best-performing algorithm. Therefore, the benefits of using ML models to predict and screen for HIV and STDs are evident.

3.8. Obstructive Sleep Apnea Syndrome

There are a number of studies that have reported the use of ML for detecting obstructive sleep apnea syndrome (OSAS). For OSAS, findings have generally been positive, as in the case of a study by Tsai et al. [62]. LR, k-nearest neighbor, CNN, naïve Bayes, RF, and SVM were all compared for screening moderate-to-severe OSAS by being trained on demographic and information-based questionnaires. The study found that BMI was the most influential parameter, and RF achieved the highest accuracy in screening for both types. In another study, Alvarez et al. trained and tested a regression SVM on polysomnography and found that the dual-channel approach was a better performer than oximetry and airflow [63]. Mencar et al. used demographic and information questionnaires again to predict OSAS severity [64]. SVM and RF were the best in classification, with the strongest average in classification being 44.7%. This study demonstrates some variability in studies attempting to define a conclusion between AI and OSAS. Overall, there is lack of literature to make a comprehensive conclusion regarding the use of ML for OSAS.

3.9. Osteoporosis

Regarding osteoporosis and related diseases, four studies have compared a number of AI models (XGBoost, LR, multiplayer perceptron, SVM, RF, ANN, extreme GB, stacking with five classifiers, and SVM with radial basis function kernel) [65–68]. Models were trained on EHR, CT and clinical data, or abdomen-pelvic CT. All studies concluded that ML methods were valid and presented great potential in screening for osteoporosis. An additional study trained ML models (RF, GB, NN, and LR) on genomic data for fracture prediction [69]. The study found that GB was the best-performing model, with an AUC score of 0.71 and an accuracy of 0.88. Ultimately, more studies are required to confirm the effectiveness of ML for predicting osteoporosis.

3.10. Chronic Conditions

Chronic obstructive pulmonary disease (COPD) is characterized by permanent lung damage and airway blockage. To enhance life quality and lower mortality rates, COPD must be diagnosed and treated early. The early identification, diagnosis, and prognosis of COPD can be aided by ML methods [70]. The likelihood of hospitalization, mortality, and COPD exacerbations have all been predicted using ML algorithms. These algorithms create predictive models using a variety of data sources, including patient demographics, clinical symptoms, and imaging data. For instance, Zeng et al. developed an ML algorithm trained on 278 candidate features [71]. The model achieved an AUROC of 0.866. Another chronic condition, chronic kidney disease (CKD), is characterized by a progressive decline in kidney function over time. Kidney failure can be prevented, and patient outcomes can be enhanced by early detection and care of CKD. The early detection, diagnosis, and management of CKD can be helped by ML algorithms. For instance, Nishat et al. developed an ML system to predict the probability of CKD. Eight supervised algorithms were developed, and RF was the best-performing mode reporting an accuracy of 99.75% [72]. At the final stage of CKD, known as ESKD, patients require dialysis or a kidney transplant. The early detection, diagnosis, and management of ESKD can be facilitated by ML algorithms. ML algorithms have been used to forecast mortality and the risk of ESKD in CKD patients. These algorithms create predictive models using a variety of data sources, including medical records, test results, and demographic information. For instance, Bai et al. trained five ML models on a longitudinal CKD cohort to predict ESKD [73]. LR, naive Bayes, and RF achieved similar predictability and sensitivity and outperformed the Kidney Failure Risk Equation. Since chronic conditions are a critical aspect of primary care, more studies involving a variety of ML models are needed to confirm MLs' potential.

3.11. Detecting COVID-19 and Influenza

ML has shown great promise in detecting and differentiating between common conditions, propagating more effective recommendations and guidelines (Figure 5). Specifically, detection research has rocketed with the rise and timeline of the COVID-19 virus [74]. Zhou et al. developed an XGBoost algorithm to distinguish between influenza and COVID-19 in case there are no laboratory results of pathogens [75]. The model used EHR data to achieve AUC scores of 0.94, 0.93, and 0.85 in the training, testing, and external validation datasets. Similarly, in Zan et al., a DL model, titled *DeepFlu*, was utilized to predict individuals at risk of symptomatic flu based on gene expression data of influenza A viruses (IAV) or the infection subtypes H1N1 or H3N2 [76]. The *DeepFlu* achieved an accuracy score of 0.70 and an AUROC of 0.787. In another study, Nadda et al. combined LSTM with an NN model to interpret patients' symptoms for disease detection [77]. For dengue and cold patients, the combination of models achieved AUCs of 0.829 and 0.776 for flu, dengue, and cold, and 0.662 for flu and cold. For influenza, Hogan et al. and Choo et al. trained multiple ML models on nasopharyngeal swab samples and the mHealth app, respectively, for influenza diagnosis and screening [78,79]. Both studies concluded that ML methods are capable of being utilized for infectious disease testing. Similar findings were presented for chronic coughs in Luo et al., where a DL model, BERT, could accurately detect chronic coughs through diagnosis and medication data [79]. Additionally, in Yoo et al., severe pharyngitis could be detected through the training of smartphone-based DL algorithms on self-taken throat images (AUROC 0.988) [80]. In summary, ML appears to be effective in screening and distinguishing between COVID-19, influenza, and related illnesses.

3.12. Detecting Atrial Fibrillation

Another large center for AI detection is atrial fibrillation (AF). Six studies have evaluated unique ways to detect AF through ML models [80–84]. Through wearable devices, countless algorithms (SVM, DNN, CNN, ENN, naive Bayesian, LR, RF, GB, and W-PPG algorithm combined with W-ECG algorithm) have been trained on primary care data, RR intervals, W-PPG and W-ECG, electrocardiogram and pulse oximetry data, or waveform

data. All studies concluded that ML is capable and has the potential to detect AF through wearable devices and through a number of different information. However, more studies to confirm these findings are required.

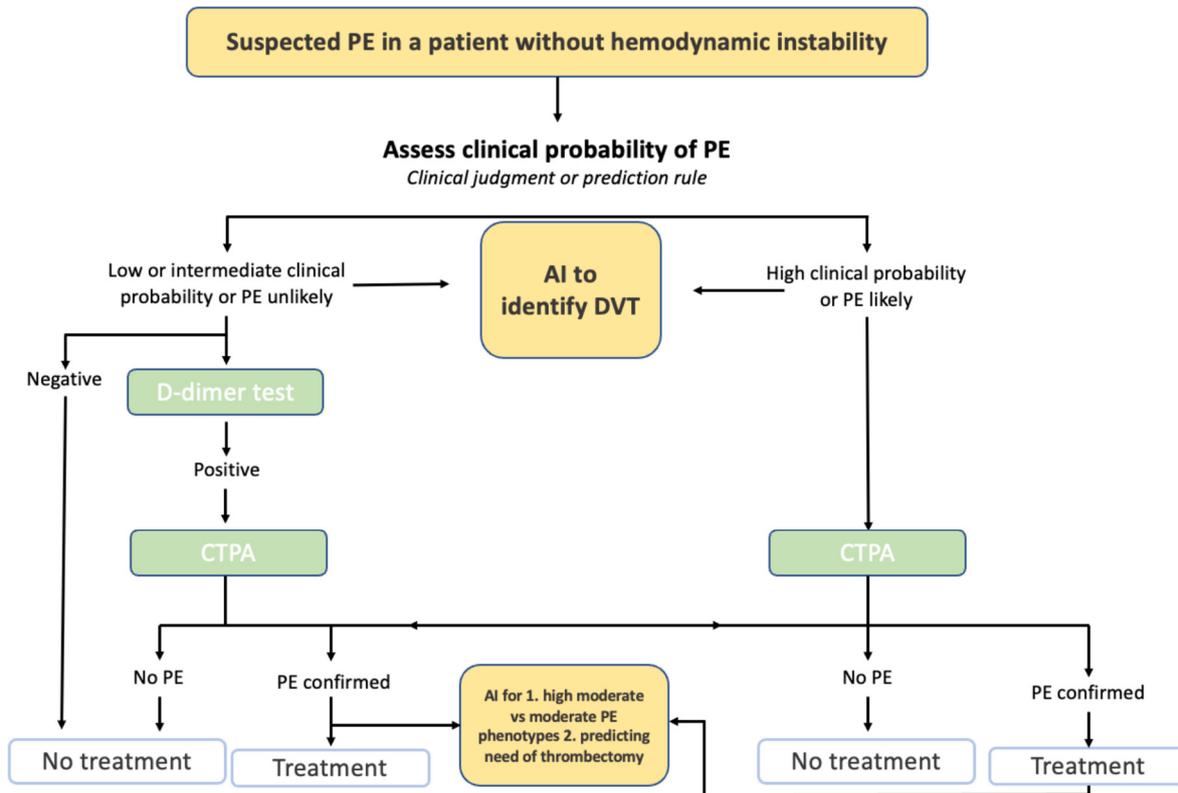


Figure 5. Example of AI in Pulmonary Embolism Evaluation. Figure Description: Current guidelines for a suspected pulmonary embolism (PE) in a patient without hemodynamic instability requires a clinical probability assessment of the PE. Based on the clinical judgment and a potential D-dimer test, a CT pulmonary angiogram is conducted to determine whether treatment or no treatment will occur. AI has the potential to be integrated into this process by potentially detecting deep vein thrombosis, detecting high moderate vs. moderate PE phenotypes, and predicting the risk of thrombectomy.

4. Limitations

While AI’s applications have been relatively positive, several limitations have set back its implementation. For one, the introduction of AI into healthcare practices raises a number of concerns, such as a lack of trust, ethical issues, and the absence of accountability [85]. Certain human traits, such as empathy, comfort, and trust, are essential to a doctor–patient relationship, and the use of AI makes these components an issue. To add on, traditionally, physicians and healthcare workers are held accountable for their practice [86]. There is no law to keep ML models intact, and there is no defined ownership to take responsibility when an AI algorithm is at fault. This drawback raises several legal and ethical concerns yet to be answered. The common novelty in ML applications across primary care requires additional clinical trials to support the potential advantages. Table 2 presents all ongoing or completed clinical trials registered in ClinicalTrials.gov and found through the keywords “Artificial Intelligence” and “Primary Care”, which were searched for ongoing or completed clinical trials investigating the role of AI in primary care. In addition, there remains mixed findings regarding the potential benefits of ML-based prediction models. For instance, in one systematic review of 71 studies, there was no evidence of a better performance from ML models compared with LR [87]. An additional drawback is that the implementation of ML is costly and would require additional education for incoming medical practitioners [88]. Regarding AI research, many studies suffer from a number of drawbacks that limit the

quality of the results. These include a small sample, retrospective data, the inability to separate pre-operative and intra-operative data, missing data, the absence of external validation, data from a single institution, and several biases.

Table 2. Clinical Trials on Artificial Intelligence in Primary Care.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT05166122	1600	Use AI to screen for diabetic retinopathy	>18 years, screened for diabetic retinopathy, with diabetes, can take retina pictures	Part of community hospital with ophthalmologist, previously diagnosed with some retinal conditions, laser retinal treatment, has other eye diseases	Recruiting
NCT05286034	4000	AI ChatBot to improve women participation in cervical cancer screening program	30–65, did not perform pap smear in last 4 years, living in deprived clusters	Outside age group, had pap smear in last 3 years, had hysterectomy including cervix, pregnant beyond 6 months, already scheduled screening appointment	Not yet recruiting
NCT04551287	16,164	Cervical cancer AI screening for cytopathological diagnosis	25–65 years old, availability of confirmed diagnosis results of cytological exam	Unsatisfactory samples of cytological exam, women diagnosed with other malignant tumors	Completed
NCT05435872	2000	AI for gastrointestinal endoscopy screening	Patients received gastroscopy and colonoscopy, endoscopic exam with AI can be accepted	Patients refusing to participate, patients with intolerance or contraindications to endoscopic exams	Recruiting
NCT05697601	2905	Finding predictors of ovarian and endometrial cancer for AI screening tool	Women with gynecological symptoms, women underwent routine gynecological exam	Unable to undergo serial gynecological exam	Recruiting
NCT04838756	100,000	AI for mammography screening	Women eligible for population-based mammography screening	None	Active, not recruiting
NCT05452993	330	AI screening for diabetic retinopathy	Adult patients with diabetes, ongoing diabetes treatment, regular pharmacy customer, informed consent	Unable to read, write, or give consent, refusing to share results with general practitioner	Not yet recruiting
NCT04778670	55,579	AI for large-scale breast screening	Participants in regular population-based breast cancer	Incomplete exam, breast implant, complete mastectomy, participant in surveillance program	Active, not recruiting
NCT05139797	300	AI-guided echo screening of rare diseases	Patients with high suspicion for cardiac amyloidosis by AI	Patients that decline to be seen at specialty clinic, patients that passed away	Recruiting
NCT05139940	2432	AI-enabled TB screening in Zambia	18 years or older with known HIV status	Individuals that do not meet inclusion criteria	Recruiting

Table 2. Cont.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT04743479	5000	AI screening of pancreatic cancer	Subject can provide informed consent, detailed questionnaire filled, and subject has one of several listed conditions	Subject has been diagnosed with pancreatic cancer or other malignant tumors in past 5 years, subject contraindicates MRI or CT, subjects is in another clinical trial	Recruiting
NCT04949776	27,000	AI for breast cancer screening	50–69 years old, women studied in the program in the set period and for the first time	Unable to give consent, breast prostheses, symptoms or signs of suspected breast cancer	Recruiting
NCT05587452	950	AI screening for colorectal cancer	Informed consent, provide blood samples, diagnosed with colorectal cancer or colorectal adenoma	Pregnant or breastfeeding, diagnosed with another cancer, selective exclusions for colorectal cancer and healthy people	Recruiting
NCT05456126	125	AI for infant motor screening	Mothers older than 20, no history of recreational drugs, married or live with fathers. Specific criteria for term and preterm infants	None	Recruiting
NCT05024591	32,714	AI for breast cancer screening	Eligible for national screening, provides consent	History or current breast cancer, currently pregnant or plans to become pregnant, history of breast surgery, has mammography for diagnostic purposes	Recruiting
NCT04732208	410	AI screening of diabetic retinopathy using smartphone camera	Over 18 years, informed consent, established cases of DM, subjects dilated for ophthalmic evaluation	Acute vision loss, contraindicated for fundus imaging, treated for retinopathy, other retinal pathologies, at risk of acute angle closure glaucoma	Completed
NCT05311046	2400	AI screening for pediatric sepsis	3 months–17 years of age, diagnosed with sepsis, blood sample collection	Participating in outside interventions, parents or LARs that do not speak English or Spanish, pregnancy	Recruiting
NCT05391659	1200	AI screening for diabetic retinopathy	Diagnosed with DM, >18 years old, informed consent, fluent in written and oral Dutch	History of diabetic retinopathy or diabetic macular edema treatment, contraindicated for imaging by fundus imaging	Recruiting
NCT04307030	5000	AI screening for congenital heart disease by heart sounds	0–18 years of age, children with or without congenital heart disease, informed consent	>18 years of age, unable to undergo echo, not able to provide informed consent	Not yet recruiting
NCT04000087	358	ECG AI-guided screening for low ejection fraction	Primary care clinicians who are part of a participating care team	Primary care clinicians working in pediatrics, acute care, nursing homes, and resident care teams	Completed

Table 2. Cont.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT04156880	1000	AI in mammography-based breast cancer screening	Women had undergone standard mammography, histopathology-proven diagnosis	Concurring lesions on mammograms, no available pathologic diagnosis or long term follow up exams, undergone breast surgery, diagnosed with other kinds of malignancy	Recruiting
NCT05645341	400	AI screening of malignant pigmented tumors on ocular surface	Dark-brown lesions on ocular surface	Non-pigmented ocular surface tumors and image quality does not meet clinical requirements	Recruiting
NCT05048095	15,500	AI in breast cancer screening	Women participating in regular breast cancer screening program	Women with breast implants or other foreign implants in mammogram and women with symptoms or signs of suspected breast cancer	Completed
NCT04894708	1572	AI for polyp detection in colonoscopy	>35 years, planned diagnostic colonoscopy's screening colonoscopy for men >50 or women >55	Colon bleeding, colon carcinoma, known polyps for removal, IBD, colonic stenosis, other suspected colon disease, follow-up care after colon cancer surgery, anticoagulant drugs, poor general condition, incomplete colonoscopy planned	Recruiting
NCT04160988	703	AI for screening diabetic retinopathy	>20 years, DM, image taken by color fundus, include includes macula and optic nerve	Color fundus image previously use, macula, optic nerve or other part is unclear	Completed
NCT04213183	1789	DL screening for hepatobiliary diseases	Quality of fundus and slit-lamp images is acceptable, more than 90% of fundus image area includes four main regions, more than 90% of slit-lamp image area includes three main regions	Images with light leakage (>10% of the area)	Completed
NCT04832594	2500	AI screening for breast cancer for supplemental MRI	Four-view screening mammography exam	Women in surveillance program, breast implants, prior breast cancer, breast feeding, MRI contraindication	Recruiting
NCT05704491	100	AI screening for diabetic retinopathy	DM diagnosis, diabetes duration >5 years, >18 years old, informed consent, fluent in writing and speaking German	History of laser treatment, contraindication to fundus imaging systems	Not yet recruiting
NCT04699864	630	AI for screening diabetic retinopathy	>18 years and older, informed consent, diagnostic for diabetes, diabetic patient followed and referred by physician	Patients less than 18 years old, no informed consent, patient already had treatment for retinal condition	Not yet recruiting

Table 2. Cont.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT04859634	2000	AI for detecting multiple ocular fundus lesions	Participants who agree to take ultra-widefield fundus images	Patients that cannot cooperate with photographer, no informed consent	Recruiting
NCT05734820	312	AI screening colonoscopy	>45 years old, referred for screening colonoscopy, adequate bowel preparation, authorized for endoscopic approach	Pregnancy, clinical condition making endoscopy invariable, history of colorectal carcinoma, IBD, no informed consent	Recruiting
NCT04859530	5886	AI smartphone for cervical cancer screening	Informed consent	No initiation of sexual intercourse, pregnancy, condition altering cervix visualization, previous hysterectomy, health not sufficient	Recruiting
NCT03773458	500	AI for large-scale screening of scoliosis	Pretreatment back photos and whole spine standing X-ray or ultrasound images	Patients considered as non-idiopathic scoliosis	Completed
NCT05704920	2722	AI for lung cancer screening	50–80 years old, active or ex-smoker, smoking history of at least 20 pack-years, informed consent, affiliated with French social security	Clinical signs of cancer, recent chest scan, health problems affecting life expectancy or limiting ability to undergo lung surgery, vulnerable people	Not yet recruiting
NCT05236855	200	AI and spectroscopy for cervical cancer screening	Women undergoing standard HPV screening	NA	Not yet recruiting
NCT05527535	34,500	AI for diabetic retinopathy screening	T1DM or T2DM, no full-time ophthalmologist, >18 years old, eligible for fundus photo imaging	T1DM or T2DM with an ophthalmologist, previous diagnosed with macular edema, history of retinal laser, other ocular disease, not eligible for fundus imaging	Not yet recruiting
NCT05745480	2	NLP for screening opioid misuse	Adults hospitalized at UW health	NA	Recruiting
NCT05490823	1000	AI smartphone for anemia screening	Informed consent	Ophthalmic or fingernail surgery in past 30 days	Recruiting
NCT04896827	244	DL and AI for DNIC	18–70 years old, chronic or no chronic pain, informed consent	CVD, Raynaud syndrome, severe psychiatric disease, injuries or loss sensitivity, pregnant women	Recruiting
NCT05752045	1389	AI for screening eye diseases	>18 years, T1DM or T2DM, presenting screening for diabetic retinopathy, benefits of social security scheme, informed consent	Patient with known DR, any condition affecting study, presenting social or psychological factors, participates in another clinical research study	Not yet recruiting

Table 2. Cont.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT05243121	5000	AI for MRI in screening breast cancer	Patients with clinical symptoms, undergoing full sequence BMRI exam, at least 6 months of follow-up results	Received therapy, contraindications of breast-enhanced MRI exams, prosthesis is implanted in affected breast, patients during lactation or pregnancy	Recruiting
NCT04996615	924	AI for identifying diabetic retinopathy and diabetic macular edema	Routine exams, routine laser treatment, diagnosed with T1DM or T2DM, presents visual acuity	Currently using AI system integrated into clinical care, inability to provide informed consent	Recruiting
NCT03975504	6000	AI for lung cancer screening	Eligible participants aged 45–75 years with one of several risk factors	Had CT scan of chest in past 12 months, history of any cancer within 5 years	Recruiting
NCT05626517	2000	Developing risk stratification tools using AI	21 years or older, sufficient English or Chinese language skills, informed consent	<21 years old, cardiac event, no informed consent	Not yet recruiting
NCT04994899	800	AI screening for mental health	13–79 years old, English-speaking	Previous participant, unable to verbally respond to standard questions, cannot participate in virtual visit, no informed consent	Recruiting
NCT05195385	2400	Lung cancer screening with low-dose CT scans	50–74 years, smoked at least 20 pack years, quit less than 15 years ago, gives consent, affiliated with social security system	Presence of clinical symptoms suggesting malignancy, evolving cancer, history of lung cancer, 2-year follow-up not possible, chest CT scan performed	Recruiting
NCT04240652	500,000	AI for diabetic retinopathy screening	T2DM or T1DM, subjects from other medical institutes are diabetes, non-diabetic patients and healthy participants	History of drug abuse, STDs, any condition not suitable for study	Recruiting
NCT04126239	1610	AI for food addiction screening test	BMI >30, able to give informed consent	Non-French speaker, unable to use internet tools	Recruiting
NCT04603404	430	Multimodality imaging in screening, diagnosis, and risk stratification of HFpEF	LVEF > 50%, NT-proBNP > 220pg/mL or BNP > pg/mL, symptoms and syndromes of HF, at least one criteria of cardiac structure	Special types of cardiomyopathies, infarction, myocardial fibrosis, severe arrhythmia, severe primary cardiac valvular disease, restrictive pericardial disease, refuses to participate in study	Recruiting
NCT05159661	1000	AI for screening brain connectivity and dementia risk estimation	Male and female 60–75 years, MCI diagnosis with MMSE > 25, MCI diagnosis with MoCa > 17	Confirmed dementia, history of cerebrovascular disease, AUD identification test, severe medical disorders associated with cognitive impairment, severe head trauma, severe mental disorders	Recruiting

Table 2. Cont.

Trial or Registry	N	Aim	Inclusion Criteria	Exclusion Criteria	Status
NCT05650086	700	AI for breast screening	Understands the study, informed consent, complies with schedule, >21 years, fits cohort specific criteria	Does not fit cohort specific criteria, unable to complete study procedures	Recruiting
NCT05426135	3000	AI for tumor risk assessment	Participants with suspected cancer, informed consent, detailed EHR data, healthy participants	Participants with primary clinical and pathological missing data, lost to follow-up, poor medical image quality	Recruiting
NCT05639348	650	AI for risk assessment of postoperative delirium	Surgical patients, >60 years old, planned postoperative hospital stay >2 days, informed consent	Preoperative delirium, insufficient knowledge in German or French, intracranial surgery, cardiac surgery, surgery within two previous weeks, unable to provide informed consent	Recruiting
NCT05466864	120	Screening of OSA using BSP	Hospitalized with acute ischemic stroke, 18–80, informed consent	History of AF, LVEF < 45%, aphasia, unstable cardiopulmonary status, recent surgery including tracheotomy in 30 days, narcotics, on O2, PAP device, ventilator, unable to understand instructions	Recruiting
NCT05655117	440	AI for detecting eye complications in diabetics	Diabetic patients aged 18–90	Severely ill patient or patient with cancer	Not yet recruiting
NCT03688906	3275	AI colorectal cancer screening test	Differs across three cohorts	Differs across three cohorts	Completed
NCT05246163	1500	AI smartphone for skin cancer detection	Patients with one or two lesions meeting one of several criteria, informed consent	Lack of informed consent	Recruiting
NCT05730192	950	AI for detection of gastrointestinal lesions in endoscopy	Screening or surveillance colonoscopy, age 40 or older, informed consent	Emergency colonoscopies, IBD, CRC, previous colonic resection, returning for elective colonoscopy, polyposis syndromes, contraindications	Not yet recruiting
NCT05566002	2000	AI evaluation of pulmonary hypertension	>18 years, previous received diagnostic imaging	Patients without RHC, quality of exams cannot meet requirement, severe loss of results	Recruiting

5. Implementing AI in Primary Care

Choosing the correct ML model for a primary care task depends on several factors, including the nature of the task, the available data, and the desired outcome (Table 3). First, a definition of the problem and the necessary data must be identified to select the appropriate model [89]. Subsequently, a suitable AI technique, such as supervised, unsupervised, or reinforcement learning, must be chosen. Following the selection of the model, evaluation of the model's performance using validation data and fine-tuning is necessary [89]. Several factors must be considered to evaluate the benefits and risks of implementing a specific AI

model into a primary care routine. Accuracy and reliability must be assessed by testing the ML model's performance on validation data [89]. Clinical relevance must be determined by evaluating whether the model is based on relevant risk factors and whether the predictions are helpful for clinical decision-making. Potential benefits such as improving patient outcomes, reducing medical errors, increasing efficiency and productivity, and enhancing the quality of care must also be assessed. Ethical implications of using the AI model in primary care, such as the responsibility of healthcare providers to explain how the AI model works and how decisions are made, and potential issues related to patient autonomy and informed consent, must be considered. Finally, the cost-effectiveness of implementing the AI model, considering the costs of development, implementation, maintenance, and training, as well as potential cost savings and benefits, must be evaluated [90]. Finally, we can anticipate a number of ML technologies, such as sophisticated chatbots and virtual assistants, decision support tools, predictive analytics, wearable technology, and population health management, to become commonplace in primary care during the next two years. These tools could aid primary care providers in making better judgements, delivering more individualized care, and spotting high-risk patients or those needing more intense interventions. However, regulatory approval, patient and healthcare provider acceptance, and integration into current clinical workflows will all be necessary before ML can be deployed. Despite these obstacles, there will likely be major advancements in integrating AI into primary care in the upcoming years, given the rate of technological advancement and the growing desire for more individualized and effective healthcare.

Table 3. Machine learning models.

ML Model	Advantages	Limitations	Clinical Applications in Primary Care
Logistic Regression	Easy to implement and interpret, handles binary and multi-class classification	Does not perform well with outliers, assumes linear relationship	Diagnostic tests, selection of treatment, prognostic modeling, predicting disease risk
Convolutional Neural Network	Excels in video and image recognition, learns hierarchical features	Needs a lot of data and resource, interpretation is limited	Image classification, diagnosing from medical imaging
Support Vector Machine	Handles non-linear decision boundaries, great generalization	Precise kernel function and hyperparameters selection, difficult with noisy data	Diagnosing disease, risk stratification, classifying clinical data
K-Nearest Neighbors	Easy, simple, handles non-linear decision boundaries	Needs a lot of memory and time, sensitivities to certain features	Assisting in disease progression through forecasting
Random Forest	Performs well with high-dimensional data, handles non-linear effects	Hard to interpret, overfits noisy data	Identifying risk factors, predicting outcomes,
Adaptive Boosting	Handles regression and classification problems, combines weak learners	Overfits with weak learners, sensitive to noisy data	Predicting risk of disease, and detecting high risk
Gradient Boosting	Performs with large datasets, handles regression and classification	Overfits with weak learners, sensitive to noisy data	Forecasting outcomes and diagnosing disease
Neural Network	Handles large datasets, performs well on speech and image recognition	Needs a lot of computational resources and data, overfits if complex	Diagnosing disease, selecting treatment, predicting risk of disease
Extreme Gradient Boosting	Fast with large datasets, handles regression and classification	Needs tuning of hyperparameters, overfits with complex weak learners	Predicting outcomes, detecting high-risk patients, diagnosing disease

Table 3. Cont.

ML Model	Advantages	Limitations	Clinical Applications in Primary Care
Decision Tree	Simple, easy, handles categorical and numerical data	Overfits with noisy data, sensitivity to variations in training	Identifying risk factors, diagnosing disease, predicting risk of disease
Deep Neural Network	Good performer with large datasets, automatically learns hierarchical features	Requires a lot of data, overfits with complex network	Diagnosing disease, detecting high-risk patients, predicting the risk of disease
Gated Recurrent Unit	Great performer with time-series data, handles variable-length sequences	Sensitivity to some conditions and parameters, poor generalization to new data	Predicting risk of disease, diagnosing diseases, and determining outcomes
XGBoost	Fast, accurate, handles regressions and classification problems	Needs tuning of hyperparameters, overfits with noisy data	Predicting outcomes, identifying risk factors
CatBoost	Handles categorical data, handles regression and classification problems	Needs resources and data, needs tuning of hyperparameters	Identifying risk factors, forecasting outcomes
Naïve Bayes	Simple, efficient, handles high-dimensional data	Independent between features, poor performer with correlated features	Diagnosing diseases, forecasting risk of disease
Logistic Model Tree	Combination of DT and LR to get non-linear effects	Overfits with noisy data, needs tuning of hyperparameters	Determining risk factors, predicting risk of disease
Long Short-Term Memory	Good performer with time-series data, handles variable-length sequence	Computational complexity, difficult interpretation, overfitting, difficult to handle long sequences	Forecasting outcomes, diagnosing diseases, forecasting risk of disease

6. Conclusions

AI in primary care and preventive medicine is a relatively new field of study that has developed endless possibilities. The applications are widespread, as seen through a number of studies on all facets of primary care. Although there is some variability within the findings of studies in specific fields, the general development and implementation of ML algorithms are successful and constructive. The models are usually more effective than previously established models or scores. Future research should focus on tackling the aforementioned limitations and furthering the research on promising sectors of primary care.

Author Contributions: Conceptualization, C.K.; methodology, C.K.; software, H.U.H.V.; validation, C.K., Z.W. and H.U.H.V.; investigation, A.H.E.-S. and C.K.; resources, C.K.; data curation, A.H.E.-S.; writing—original draft preparation, A.H.E.-S.; writing—review and editing, A.H.E.-S., H.U.H.V., Z.W., C.K. and B.S.G.; visualization, H.U.H.V.; supervision, C.K.; project administration. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Collins, C.; Dennehy, D.; Conboy, K.; Mikalef, P. Artificial intelligence in information systems research: A systematic literature review and research agenda. *Int. J. Inf. Manag.* **2021**, *60*, 102383. [[CrossRef](#)]
2. Kersting, K. Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines. *Front. Big Data* **2018**, *1*, 6. [[CrossRef](#)] [[PubMed](#)]
3. Ghuwalewala, S.; Kulkarni, V.; Pant, R.; Kharat, A. Levels of Autonomous Radiology. *Interact. J. Med. Res.* **2022**, *11*, e38655. [[CrossRef](#)]
4. Bignami, E.G.; Cozzani, F.; Del Rio, P.; Bellini, V. Artificial intelligence and perioperative medicine. *Minerva Anesthesiol.* **2020**, *87*, 755–756. [[CrossRef](#)] [[PubMed](#)]
5. Chiew, C.J.; Liu, N.; Wong, T.H.; Sim, Y.E.; Abdullah, H.R. Utilizing Machine Learning Methods for Preoperative Prediction of Postsurgical Mortality and Intensive Care Unit Admission. *Ann. Surg.* **2020**, *272*, 1133–1139. [[CrossRef](#)]
6. Fernandes, M.P.B.; de la Hoz, M.A.; Rangasamy, V.; Subramaniam, B. Machine Learning Models with Preoperative Risk Factors and Intraoperative Hypotension Parameters Predict Mortality After Cardiac Surgery. *J. Cardiothorac. Vasc. Anesth.* **2021**, *35*, 857–865. [[CrossRef](#)] [[PubMed](#)]
7. Jalali, A.; Lonsdale, H.; Do, N.; Peck, J.; Gupta, M.; Kutty, S.; Ghazarian, S.R.; Jacobs, J.P.; Rehman, M.; Ahumada, L.M. Deep Learning for Improved Risk Prediction in Surgical Outcomes. *Sci. Rep.* **2020**, *10*, 9289. [[CrossRef](#)]
8. Pfitzner, B.; Chromik, J.; Brabender, R.; Fischer, E.; Kromer, A.; Winter, A.; Moosburner, S.; Sauer, I.M.; Malinka, T.; Pratschke, J.; et al. Perioperative Risk Assessment in Pancreatic Surgery Using Machine Learning. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Scotland, UK, 1–5 November 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 2211–2214. [[CrossRef](#)]
9. Sahara, K.; Paredes, A.Z.; Tsilimigras, D.I.; Sasaki, K.; Moro, A.; Hyer, J.M.; Mehta, R.; Farooq, S.A.; Wu, L.; Endo, I.; et al. Machine learning predicts unpredicted deaths with high accuracy following hepatopancreatic surgery. *HepatoBiliary Surg. Nutr.* **2021**, *10*, 20–30. [[CrossRef](#)]
10. COVIDSurg Collaborativ; Dajti, I.; Valenzuela, J.I.; Boccalatte, L.A.; Gemelli, N.A.; Smith, D.E.; Dudi-Venkata, N.N.; Kroon, H.M.; Sammour, T.; Roberts, M.; et al. Machine learning risk prediction of mortality for patients undergoing surgery with perioperative SARS-CoV-2: The COVIDSurg mortality score. *Br. J. Surg.* **2021**, *108*, 1274–1292. [[CrossRef](#)]
11. Xue, B.; Li, D.; Lu, C.; King, C.R.; Wildes, T.; Avidan, M.S.; Kannampallil, T.; Abraham, J. Use of Machine Learning to Develop and Evaluate Models Using Preoperative and Intraoperative Data to Identify Risks of Postoperative Complications. *JAMA Netw. Open* **2021**, *4*, e212240. [[CrossRef](#)]
12. Corey, K.M.; Kashyap, S.; Lorenzi, E.; Lagoo-Deenadayalan, S.A.; Heller, K.; Whalen, K.; Balu, S.; Heflin, M.T.; McDonald, S.R.; Swaminathan, M.; et al. Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): A retrospective, single-site study. *PLOS Med.* **2018**, *15*, e1002701. [[CrossRef](#)]
13. Bonde, A.; Varadarajan, K.M.; Bonde, N.; Troelsen, A.; Muratoglu, O.K.; Malchau, H.; Yang, A.D.; Alam, H.; Sillesen, M. Assessing the utility of deep neural networks in predicting postoperative surgical complications: A retrospective study. *Lancet Digit. Health* **2021**, *3*, e471–e485. [[CrossRef](#)]
14. Zhao, H.; Zhang, X.; Xu, Y.; Gao, L.; Ma, Z.; Sun, Y.; Wang, W. Predicting the Risk of Hypertension Based on Several Easy-to-Collect Risk Factors: A Machine Learning Method. *Front. Public Health* **2021**, *9*, 619429. [[CrossRef](#)]
15. Alkaabi, L.A.; Ahmed, L.S.; Al Attiyah, M.F.; Abdel-Rahman, M.E. Predicting hypertension using machine learning: Findings from Qatar Biobank Study. *PLoS ONE* **2020**, *15*, e0240370. [[CrossRef](#)]
16. Ye, C.; Fu, T.; Hao, S.; Zhang, Y.; Wang, O.; Jin, B.; Xia, M.; Liu, M.; Zhou, X.; Wu, Q.; et al. Prediction of Incident Hypertension Within the Next Year: Prospective Study Using Statewide Electronic Health Records and Machine Learning. *J. Med. Internet Res.* **2018**, *20*, e22. [[CrossRef](#)]
17. LaFreniere, D.; Zulkernine, F.; Barber, D.; Martin, K. Using machine learning to predict hypertension from a clinical dataset. In Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 6–9 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–7. [[CrossRef](#)]
18. Khalid, S.G.; Zhang, J.; Chen, F.; Zheng, D. Blood Pressure Estimation Using Photoplethysmography Only: Comparison between Different Machine Learning Approaches. *J. Healthc. Eng.* **2018**, *2018*, 1548647. [[CrossRef](#)]
19. Myers, K.D.; Knowles, J.W.; Staszak, D.; Shapiro, M.D.; Howard, W.; Yadava, M.; Zuzick, D.; Williamson, L.; Shah, N.H.; Banda, J.; et al. Precision screening for familial hypercholesterolaemia: A machine learning study applied to electronic health encounter data. *Lancet Digit. Health* **2019**, *1*, e393–e402. [[CrossRef](#)]
20. Pina, A.; Helgadottir, S.; Mancina, R.M.; Pavanello, C.; Pirazzi, C.; Montalcini, T.; Henriques, R.; Calabresi, L.; Wiklund, O.; Macedo, M.P.; et al. Virtual genetic diagnosis for familial hypercholesterolemia powered by machine learning. *Eur. J. Prev. Cardiol.* **2020**, *27*, 1639–1646. [[CrossRef](#)]
21. Liu, Y.; Zhang, Q.; Zhao, G.; Liu, G.; Liu, Z. Deep Learning-Based Method of Diagnosing Hyperlipidemia and Providing Diagnostic Markers Automatically. *Diabetes Metab. Syndr. Obes. Targets Ther.* **2020**, *13*, 679–691. [[CrossRef](#)]
22. Tsigalou, C.; Panopoulou, M.; Papadopoulos, C.; Karvelas, A.; Tsairidis, D.; Anagnostopoulos, K. Estimation of low-density lipoprotein cholesterol by machine learning methods. *Clin. Chim. Acta* **2021**, *517*, 108–116. [[CrossRef](#)]

23. Çubukçu, H.C.; Topcu, D.I. Estimation of Low-Density Lipoprotein Cholesterol Concentration Using Machine Learning. *Lab. Med.* **2022**, *53*, 161–171. [[CrossRef](#)] [[PubMed](#)]
24. Weng, S.F.; Reys, J.; Kai, J.; Garibaldi, J.M.; Qureshi, N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE* **2017**, *12*, e0174944. [[CrossRef](#)] [[PubMed](#)]
25. Zhao, J.; Feng, Q.; Wu, P.; Lupu, R.A.; Wilke, R.A.; Wells, Q.S.; Denny, J.C.; Wei, W.-Q. Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Sci. Rep.* **2019**, *9*, 717. [[CrossRef](#)] [[PubMed](#)]
26. Kusunose, K.; Hirata, Y.; Tsuji, T.; Kotoku, J.; Sata, M. Deep learning to predict elevated pulmonary artery pressure in patients with suspected pulmonary hypertension using standard chest X-ray. *Sci. Rep.* **2020**, *10*, 19311. [[CrossRef](#)]
27. Madani, A.; Moradi, M.; Karargyris, A.; Syeda-Mahmood, T. Chest X-ray generation and data augmentation for cardiovascular abnormality classification. In *Medical Imaging 2018: Image Processing*; Angelini, E.D., Landman, B.A., Eds.; SPIE: Houston, TX, USA, 2018; p. 57. [[CrossRef](#)]
28. Ambale-Venkatesh, B.; Yang, X.; Wu, C.O.; Liu, K.; Hundley, W.G.; McClelland, R.; Gomes, A.S.; Folsom, A.R.; Shea, S.; Guallar, E.; et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. *Circ. Res.* **2017**, *121*, 1092–1101. [[CrossRef](#)]
29. Alaa, A.M.; Bolton, T.; Di Angelantonio, E.; Rudd, J.H.F.; van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS ONE* **2019**, *14*, e0213653. [[CrossRef](#)]
30. Pfohl, S.; Marafino, B.; Coulet, A.; Rodriguez, F.; Palaniappan, L.; Shah, N.H. Creating Fair Models of Atherosclerotic Cardiovascular Disease Risk. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, Honolulu, HI, USA, 27–28 January 2019. [[CrossRef](#)]
31. Al'aref, S.J.; Maliakal, G.; Singh, G.; van Rosendaal, A.R.; Ma, X.; Xu, Z.; Alawamlh, O.A.H.; Lee, B.; Pandey, M.; Achenbach, S.; et al. Machine learning of clinical variables and coronary artery calcium scoring for the prediction of obstructive coronary artery disease on coronary computed tomography angiography: Analysis from the CONFIRM registry. *Eur. Heart J.* **2020**, *41*, 359–367. [[CrossRef](#)]
32. Ting, D.S.W.; Cheung, C.Y.-L.; Lim, G.; Tan, G.S.W.; Quang, N.D.; Gan, A.; Hamzah, H.; Garcia-Franco, R.; Yeo, I.Y.S.; Lee, S.Y.; et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* **2017**, *318*, 2211–2223. [[CrossRef](#)]
33. Poplin, R.; Varadarajan, A.V.; Blumer, K.; Liu, Y.; McConnell, M.V.; Corrado, G.S.; Peng, L.; Webster, D.R. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat. Biomed. Eng.* **2018**, *2*, 158–164. [[CrossRef](#)]
34. Kim, Y.D.; Noh, K.J.; Byun, S.J.; Lee, S.; Kim, T.; Sunwoo, L.; Lee, K.J.; Kang, S.-H.; Park, K.H.; Park, S.J. Effects of Hypertension, Diabetes, and Smoking on Age and Sex Prediction from Retinal Fundus Images. *Sci. Rep.* **2020**, *10*, 4623. [[CrossRef](#)]
35. Ravaut, M.; Harish, V.; Sadeghi, H.; Leung, K.K.; Volkovs, M.; Kornas, K.; Watson, T.; Poutanen, T.; Rosella, L.C. Development and Validation of a Machine Learning Model Using Administrative Health Data to Predict Onset of Type 2 Diabetes. *JAMA Netw. Open* **2021**, *4*, e2111315. [[CrossRef](#)]
36. Ravaut, M.; Sadeghi, H.; Leung, K.K.; Volkovs, M.; Kornas, K.; Harish, V.; Watson, T.; Lewis, G.F.; Weisman, A.; Poutanen, T.; et al. Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data. *Npj Digit. Med.* **2021**, *4*, 24. [[CrossRef](#)]
37. Deberneh, H.M.; Kim, I. Prediction of Type 2 Diabetes Based on Machine Learning Algorithm. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3317. [[CrossRef](#)]
38. Alhassan, Z.; McGough, A.S.; Alshammari, R.; Daghestani, T.; Budgen, D.; Al Moubayed, N. Type-2 Diabetes Mellitus Diagnosis from Time Series Clinical Data Using Deep Learning Models. In *Artificial Neural Networks and Machine Learning—ICANN 2018*; Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; Volume 11141, pp. 468–478. [[CrossRef](#)]
39. Boutilier, J.J.; Chan, T.C.Y.; Ranjan, M.; Deo, S. Risk Stratification for Early Detection of Diabetes and Hypertension in Resource-Limited Settings: Machine Learning Analysis. *J. Med. Internet Res.* **2021**, *23*, e20123. [[CrossRef](#)]
40. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
41. Alharbi, F.; Vakanski, A. Machine Learning Methods for Cancer Classification Using Gene Expression Data: A Review. *Bioengineering* **2023**, *10*, 173. [[CrossRef](#)]
42. Ardila, D.; Kiraly, A.P.; Bharadwaj, S.; Choi, B.; Reicher, J.J.; Peng, L.; Tse, D.; Etemadi, M.; Ye, W.; Corrado, G.; et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **2019**, *25*, 954–961. [[CrossRef](#)]
43. Gould, M.K.; Huang, B.Z.; Tammemagi, M.C.; Kinar, Y.; Shiff, R. Machine Learning for Early Lung Cancer Identification Using Routine Clinical and Laboratory Data. *Am. J. Respir. Crit. Care Med.* **2021**, *204*, 445–453. [[CrossRef](#)]
44. Yeh, M.C.-H.; Wang, Y.-H.; Yang, H.-C.; Bai, K.-J.; Wang, H.-H.; Li, Y.-C.J. Artificial Intelligence-Based Prediction of Lung Cancer Risk Using Nonimaging Electronic Medical Records: Deep Learning Approach. *J. Med. Internet Res.* **2021**, *23*, e26256. [[CrossRef](#)]
45. Guo, Y.; Yin, S.; Chen, S.; Ge, Y. Predictors of underutilization of lung cancer screening: A machine learning approach. *Eur. J. Cancer Prev.* **2022**, *31*, 523–529. [[CrossRef](#)]

46. Mehmood, M.; Rizwan, M.; MI, M.G.; Abbas, S. Machine Learning Assisted Cervical Cancer Detection. *Front. Public Health* **2021**, *9*, 788376. [[CrossRef](#)] [[PubMed](#)]
47. Rahaman, M.; Li, C.; Yao, Y.; Kulwa, F.; Wu, X.; Li, X.; Wang, Q. DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. *Comput. Biol. Med.* **2021**, *136*, 104649. [[CrossRef](#)] [[PubMed](#)]
48. Wentzensen, N.; Lahrmann, B.; Clarke, M.A.; Kinney, W.; Tokugawa, D.; Poitras, N.; Locke, A.; Bartels, L.; Krauthoff, A.; Walker, J.; et al. Accuracy and Efficiency of Deep-Learning-Based Automation of Dual Stain Cytology in Cervical Cancer Screening. *JNCI J. Natl. Cancer Inst.* **2020**, *113*, 72–79. [[CrossRef](#)] [[PubMed](#)]
49. Shen, L.; Margolies, L.R.; Rothstein, J.H.; Fluder, E.; McBride, R.; Sieh, W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Sci. Rep.* **2019**, *9*, 12495. [[CrossRef](#)]
50. Buda, M.; Saha, A.; Walsh, R.; Ghatge, S.; Li, N.; Swiecicki, A.; Lo, J.Y.; Mazurowski, M.A. A Data Set and Deep Learning Algorithm for the Detection of Masses and Architectural Distortions in Digital Breast Tomosynthesis Images. *JAMA Netw. Open* **2021**, *4*, e2119100. [[CrossRef](#)]
51. Maghsoudi, O.H.; Gastounioti, A.; Scott, C.; Pantalone, L.; Wu, F.-F.; Cohen, E.A.; Winham, S.; Conant, E.F.; Vachon, C.; Kontos, D. Deep-LIBRA: An artificial-intelligence method for robust quantification of breast density with independent validation in breast cancer risk assessment. *Med. Image Anal.* **2021**, *73*, 102138. [[CrossRef](#)]
52. Ming, C.; Viassolo, V.; Probst-Hensch, N.; Dinov, I.D.; Chappuis, P.O.; Katapodi, M.C. Machine learning-based lifetime breast cancer risk reclassification compared with the BOADICEA model: Impact on screening recommendations. *Br. J. Cancer* **2020**, *123*, 860–867. [[CrossRef](#)]
53. Perera, M.; Mirchandani, R.; Papa, N.; Breemer, G.; Effeindzourou, A.; Smith, L.; Swindle, P.; Smith, E. PSA-based machine learning model improves prostate cancer risk stratification in a screening population. *World J. Urol.* **2021**, *39*, 1897–1902. [[CrossRef](#)]
54. Chiu, P.K.-F.; Shen, X.; Wang, G.; Ho, C.-L.; Leung, C.-H.; Ng, C.-F.; Choi, K.-S.; Teoh, J.Y.-C. Enhancement of prostate cancer diagnosis by machine learning techniques: An algorithm development and validation study. *Prostate Cancer Prostatic Dis.* **2021**, *25*, 672–676. [[CrossRef](#)]
55. Beinecke, J.M.; Anders, P.; Schurrat, T.; Heider, D.; Luster, M.; Librizzi, D.; Hauschild, A.-C. Evaluation of machine learning strategies for imaging confirmed prostate cancer recurrence prediction on electronic health records. *Comput. Biol. Med.* **2022**, *143*, 105263. [[CrossRef](#)]
56. Turbé, V.; Herbst, C.; Mngomezulu, T.; Meshkinfamfard, S.; Dlamini, N.; Mhlongo, T.; Smit, T.; Cherepanova, V.; Shimada, K.; Budd, J.; et al. Deep learning of HIV field-based rapid tests. *Nat. Med.* **2021**, *27*, 1165–1170. [[CrossRef](#)]
57. Bao, Y.; Medland, N.A.; Fairley, C.K.; Wu, J.; Shang, X.; Chow, E.P.; Xu, X.; Ge, Z.; Zhuang, X.; Zhang, L. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *J. Infect.* **2020**, *82*, 48–59. [[CrossRef](#)]
58. Marcus, J.L.; Hurley, L.B.; Krakower, D.S.; Alexeeff, S.; Silverberg, M.J.; Volk, J.E. Use of electronic health record data and machine learning to identify candidates for HIV pre-exposure prophylaxis: A modelling study. *Lancet HIV* **2019**, *6*, e688–e695. [[CrossRef](#)]
59. Elder, H.R.; Gruber, S.; Willis, S.J.; Cocoros, N.; Callahan, M.; Flagg, E.W.; Klompas, M.; Hsu, K.K. Can Machine Learning Help Identify Patients at Risk for Recurrent Sexually Transmitted Infections? *Sex. Transm. Dis.* **2020**, *48*, 56–62. [[CrossRef](#)]
60. Gadalla, A.A.H.; Friberg, I.M.; Kift-Morgan, A.; Zhang, J.; Eberl, M.; Topley, N.; Weeks, I.; Cuff, S.; Wootton, M.; Gal, M.; et al. Identification of clinical and urine biomarkers for uncomplicated urinary tract infection using machine learning algorithms. *Sci. Rep.* **2019**, *9*, 19694. [[CrossRef](#)]
61. Taylor, R.A.; Moore, C.L.; Cheung, K.-H.; Brandt, C. Predicting urinary tract infections in the emergency department with machine learning. *PLoS ONE* **2018**, *13*, e0194085. [[CrossRef](#)]
62. Tsai, C.-Y.; Liu, W.-T.; Lin, Y.-T.; Lin, S.-Y.; Houghton, R.; Hsu, W.-H.; Wu, D.; Lee, H.-C.; Wu, C.-J.; Li, L.Y.J.; et al. Machine learning approaches for screening the risk of obstructive sleep apnea in the Taiwan population based on body profile. *Inform. Health Soc. Care* **2021**, *47*, 373–388. [[CrossRef](#)]
63. Álvarez, D.; Cerezo-Hernández, A.; Crespo, A.; Gutiérrez-Tobal, G.C.; Vaquerizo-Villar, F.; Barroso-García, V.; Moreno, F.; Arroyo, C.A.; Ruiz, T.; Hornero, R.; et al. A machine learning-based test for adult sleep apnoea screening at home using oximetry and airflow. *Sci. Rep.* **2020**, *10*, 5332. [[CrossRef](#)]
64. Mencar, C.; Gallo, C.; Mantero, M.; Tarsia, P.; Carpagnano, G.E.; Barbaro, M.P.F.; Lacedonia, D. Application of machine learning to predict obstructive sleep apnea syndrome severity. *Health Inform. J.* **2020**, *26*, 298–317. [[CrossRef](#)]
65. Park, H.W.; Jung, H.; Back, K.Y.; Choi, H.J.; Ryu, K.S.; Cha, H.S.; Lee, E.K.; Hong, A.R.; Hwangbo, Y. Application of Machine Learning to Identify Clinically Meaningful Risk Group for Osteoporosis in Individuals Under the Recommended Age for Dual-Energy X-Ray Absorptiometry. *Calcif. Tissue Int.* **2021**, *109*, 645–655. [[CrossRef](#)]
66. Kim, S.K.; Yoo, T.K.; Oh, E.; Kim, D.W. Osteoporosis risk prediction using machine learning and conventional methods. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 188–191. [[CrossRef](#)]
67. Liu, L.; Si, M.; Ma, H.; Cong, M.; Xu, Q.; Sun, Q.; Wu, W.; Wang, C.; Fagan, M.J.; Mur, L.A.J.; et al. A hierarchical opportunistic screening model for osteoporosis using machine learning applied to clinical data and CT images. *BMC Bioinform.* **2022**, *23*, 63. [[CrossRef](#)] [[PubMed](#)]

68. Lim, H.K.; Ha, H.I.; Park, S.-Y.; Han, J. Prediction of femoral osteoporosis using machine-learning analysis with radiomics features and abdomen-pelvic CT: A retrospective single center preliminary study. *PLoS ONE* **2021**, *16*, e0247330. [[CrossRef](#)] [[PubMed](#)]
69. Wu, Q.; Nasoz, F.; Jung, J.; Bhattarai, B.; Han, M.V. Machine Learning Approaches for Fracture Risk Assessment: A Comparative Analysis of Genomic and Phenotypic Data in 5130 Older Men. *Calcif. Tissue Int.* **2020**, *107*, 353–361. [[CrossRef](#)] [[PubMed](#)]
70. Moslemi, A.; Kontogianni, K.; Brock, J.; Wood, S.; Herth, F.; Kirby, M. Differentiating COPD and asthma using quantitative CT imaging and machine learning. *Eur. Respir. J.* **2022**, *60*, 2103078. [[CrossRef](#)] [[PubMed](#)]
71. Zeng, S.; Arjomandi, M.; Tong, Y.; Liao, Z.C.; Luo, G. Developing a Machine Learning Model to Predict Severe Chronic Obstructive Pulmonary Disease Exacerbations: Retrospective Cohort Study. *J. Med. Internet Res.* **2022**, *24*, e28953. [[CrossRef](#)]
72. Nishat, M.; Faisal, F.; Dip, R.; Nasrullah, S.; Ahsan, R.; Shikder, F.; Asif, M.A. Hoque A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms. *EAI Endorsed Trans. Pervasive Health Technol.* **2018**, *7*, 170671. [[CrossRef](#)]
73. Bai, Q.; Su, C.; Tang, W.; Li, Y. Machine learning to predict end stage kidney disease in chronic kidney disease. *Sci. Rep.* **2022**, *12*, 8377. [[CrossRef](#)]
74. Heidari, A.; Navimipour, N.J.; Unal, M.; Toumaj, S. Machine learning applications for COVID-19 outbreak management. *Neural Comput. Appl.* **2022**, *34*, 15313–15348. [[CrossRef](#)]
75. Zhou, X.; Wang, Z.; Li, S.; Liu, T.; Wang, X.; Xia, J.; Zhao, Y. Machine Learning-Based Decision Model to Distinguish Between COVID-19 and Influenza: A Retrospective, Two-Centered, Diagnostic Study. *Risk Manag. Healthc. Policy* **2021**, *14*, 595–604. [[CrossRef](#)]
76. Zan, A.; Xie, Z.-R.; Hsu, Y.-C.; Chen, Y.-H.; Lin, T.-H.; Chang, Y.-S.; Chang, K.Y. DeepFlu: A deep learning approach for forecasting symptomatic influenza A infection based on pre-exposure gene expression. *Comput. Methods Programs Biomed.* **2021**, *213*, 106495. [[CrossRef](#)]
77. Nadda, W.; Boonchieng, W.; Boonchieng, E. Influenza, dengue and common cold detection using LSTM with fully connected neural network and keywords selection. *BioData Min.* **2022**, *15*, 5. [[CrossRef](#)]
78. Hogan, C.A.; Rajpurkar, P.; Sowrirajan, H.; Phillips, N.A.; Le, A.T.; Wu, M.; Garamani, N.; Sahoo, M.K.; Wood, M.L.; Huang, C.; et al. Nasopharyngeal metabolomics and machine learning approach for the diagnosis of influenza. *EbioMedicine* **2021**, *71*, 103546. [[CrossRef](#)]
79. Choo, H.; Kim, M.; Choi, J.; Shin, J.; Shin, S.-Y. Influenza Screening via Deep Learning Using a Combination of Epidemiological and Patient-Generated Health Data: Development and Validation Study. *J. Med. Internet Res.* **2020**, *22*, e21369. [[CrossRef](#)]
80. Lown, M.; Brown, M.; Brown, C.; Yue, A.M.; Shah, B.N.; Corbett, S.J.; Lewith, G.; Stuart, B.; Moore, M.; Little, P. Machine learning detection of Atrial Fibrillation using wearable technology. *PLoS ONE* **2020**, *15*, e0227401. [[CrossRef](#)]
81. Ali, F.; Hasan, B.; Ahmad, H.; Hoodbhoy, Z.; Bhuriwala, Z.; Hanif, M.; Ansari, S.U.; Chowdhury, D. Detection of subclinical rheumatic heart disease in children using a deep learning algorithm on digital stethoscope: A study protocol. *BMJ Open* **2021**, *11*, e044070. [[CrossRef](#)]
82. Kwon, S.; Hong, J.; Choi, E.-K.; Lee, E.; Hostallero, D.E.; Kang, W.J.; Lee, B.; Jeong, E.-R.; Koo, B.-K.; Oh, S.; et al. Deep Learning Approaches to Detect Atrial Fibrillation Using Photoplethysmographic Signals: Algorithms Development Study. *JMIR mHealth uHealth* **2019**, *7*, e12770. [[CrossRef](#)]
83. Tiwari, P.; Colborn, K.L.; Smith, D.E.; Xing, F.; Ghosh, D.; Rosenberg, M.A. Assessment of a Machine Learning Model Applied to Harmonized Electronic Health Record Data for the Prediction of Incident Atrial Fibrillation. *JAMA Netw. Open* **2020**, *3*, e1919396. [[CrossRef](#)]
84. Sekelj, S.; Sandler, B.; Johnston, E.; Pollock, K.G.; Hill, N.R.; Gordon, J.; Tsang, C.; Khan, S.; Ng, F.S.; Farooqui, U. Detecting undiagnosed atrial fibrillation in UK primary care: Validation of a machine learning prediction algorithm in a retrospective cohort study. *Eur. J. Prev. Cardiol.* **2021**, *28*, 598–605. [[CrossRef](#)]
85. Kelly, C.J.; Karthikesalingam, A.; Suleyman, M.; Corrado, G.; King, D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* **2019**, *17*, 195. [[CrossRef](#)]
86. Sunarti, S.; Rahman, F.F.; Naufal, M.; Risky, M.; Febriyanto, K.; Masnina, R. Artificial intelligence in healthcare: Opportunities and risk for future. *Gac. Sanit.* **2021**, *35*, S67–S70. [[CrossRef](#)]
87. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **2019**, *110*, 12–22. [[CrossRef](#)]
88. Ahuja, A.S. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* **2019**, *7*, e7702. [[CrossRef](#)] [[PubMed](#)]
89. Raschka, S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *arXiv* **2018**, arXiv:1811.12808. [[CrossRef](#)]
90. de Vos, J.; Visser, L.A.; de Beer, A.A.; Fornasa, M.; Thorall, P.J.; Elbers, P.W.; Cinà, G. The Potential Cost-Effectiveness of a Machine Learning Tool That Can Prevent Untimely Intensive Care Unit Discharge. *Value Health* **2021**, *25*, 359–367. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.