Systematic Reviews of Machine Learning in Healthcare: A Literature Review

katarzyna kolasa (kkolasa@kozminski.edu.pl)
Kozminski University
Bisrat Yeshewas Admassu
Kozminski University
Malwina Hołownia
Kozminski University
Katarzyna Kędzior
Kozminski University
Jean-Etienne Poirrier
PAREXEL International (United Kingdom)
Stefano Perni
PAREXEL International (United Kingdom)

Research Article

Keywords:

Posted Date: June 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3035548/v1

Abstract

The increasing availability of data and computing power has made machine learning (ML) a viable approach to faster, more efficient healthcare delivery. To exploit the potential of data-driven technologies, further integration of artificial intelligence (AI) into healthcare systems is warranted. A systematic literature review (SLR) of published SLRs evaluated evidence of ML applications in healthcare settings published in PubMed, IEEE Xplore, Scopus, Web of Science, EBSCO, and the Cochrane Library up to March 2023. Studies were classified based on the disease area and the type of ML algorithm used. In total, 220 SLRs covering 10,462 ML algorithms were identified, the majority of which aimed at solutions towards clinical prediction, categorisation, and disease prognosis in oncology and neurology primarily using imaging data. Accuracy, specificity, and sensitivity were 56%, 28%, and 25%, respectively. Internal validation was reported in 53% of the ML algorithms and external validation in below 1%. The most common modelling approach was neural networks (2,454 ML algorithms), followed by support vector machine and random forest/decision trees (1,578 and 1,522 ML algorithms, respectively). The review indicated that there is potential for greater adoption of AI in healthcare, with 10,462 ML algorithms identified compared to 523 approved by the Food and Drug Administration (FDA). However, the considerable reporting gaps call for more effort towards internal and external validation. Greater accessibility to healthcare data for developers can ensure the faster adoption of ML algorithms.

Introduction

Along with other sectors, medicine has become a prominent beneficiary of artificial intelligence (AI)-driven innovations, owing to the growing availability of data. The transformation of healthcare began with the widespread adoption of electronic health records (EHRs) in the early 1990s, with up to 93% of primary care doctors using EHR across 24 OECD countries in 2021 (1).

The growing number of new data sources such as sensors, wearables, and mobile applications is transforming healthcare. The digital footprint of a patient’s journey produces new insights that inform decision-making processes and makes them readily available for developing machine learning (ML) algorithms.

Therefore, the abundance of data can help healthcare organisations develop a more comprehensive picture of a patient's health over time and also introduce efficiency gains with new data-driven technologies.

As AI's potential to improve health outcomes and increase physician efficiency has been realised, the tech industry is a key player, alongside pharma and medtech, in the healthcare sector.

Several experts claim that medicine is already moving from the past decade that focused on ML development to the subsequent decade, driven by the challenges of ensuring ML algorithm deployment in clinical settings (2).

Although the International Medical Device Regulators Forum (IMDRF) introduced the terms “software as a medical device” (SaMD) and “software in a medical device” in 2013, there have been limited efforts thus far to introduce ML algorithms in healthcare financing system similarly as medical devices and pharmaceuticals.

To understand the state-of-the-art regarding the availability of AI solutions in healthcare, we conducted a review of systematic literature reviews (SLRs) covering ML algorithms developed for medical purposes. The objectives of our research were twofold: First, to describe the number of ML solutions already available in healthcare. Second, we assessed the types of data commonly reported in scientific publications on ML algorithms. Based on our review, we recommend actions for developers and healthcare payers to facilitate AI integration into medicine.

Methodology

This review was unregistered, and a formal protocol was not prepared. Following the inclusion criteria “SRL” and “ML” were used with additional key terms such as outcome prediction, diagnosis, screening and/or treatment of any disease. Only publications reporting the use of ML in healthcare, written in English and published in peer-reviewed journals until 27 March 2023 were included. Searches were conducted in PubMed (Pubmed. ncbi. nlm. nih. gov), IEEE Xplore (ieeexplore.ieee.org), Scopus (www.scopus.com), Web of Science (www.webofknowledge.com), EBSCO (www.ebsco.com), and the Cochrane Library (www.cochranelibrary.com).

Four researchers (hereafter referred to as reviewers) performed the initial review in three steps.

1. Identification: The titles, keywords, and abstracts of all identified publications were independently screened for relevance by two reviewers. Each reviewer decided whether the publication was relevant for further review, and consensus resulted in inclusion or exclusion. Any remaining conflict was resolved through dialogue.
2. Full-text screening: The full texts of all publications proceeding to this step were obtained and read independently by two reviewers. Each reviewer decided whether a publication was relevant for inclusion in the review or for data extraction. A consensus resulted in inclusion or exclusion at this step and any remaining conflicts were resolved through dialogue.
3. Data extraction: Data were extracted from each SLR in two phases. First, two recent checklists were analysed to define the set of review criteria (3, 4). Second, a random sample of 30 SLRs was analysed to assess the most commonly reported information across the included publications.

All identified publications were entered into the Covidence systematic review software for the remainder of the review. The PRISMA guidelines for reporting systematic reviews were followed (5).

Data extraction was initiated after the initial process. Each SLR was reviewed for basic descriptive statistics, including quality assessment and reporting methods, along with an assignment to one of the three categories.

1. Categorization (classification of data into categories or clusters)
2. Prediction (making predictions regarding outputs providing historical data)
3. Discovery (analysis of the structure of data)

ICD-10 codes were used to analyse the therapeutic area covered by the SLRs, and basic statistics such as the sources of data, accuracy, specificity, and sensitivity were extracted from the included SLRs for each publication included separately along with the methods of validation and handling missing data. Details regarding the external validation with respect to the comparison of AI against humans were also extracted and reviewed, not only from systematic literature reviews but also from primary studies. The details of the types of ML techniques were also extracted, and the number of primary studies reporting the use of different ML algorithm typologies was determined for each SLR included.

Results

A total of 2,342 SLRs were identified. Based on the title and abstract reviews, 1,233 hits were removed during the identification phase. A total of 686 duplicates were identified (Figure 1). The screening phase included 423 publications. After full-text analysis, 220 articles (6-226) covering 10,462 ML algorithms were finally included in the review (Figure 1).

The number of studies covered by each SLR varied from 4 (166) to 921 (83) articles (Table 1). Approximately 88% of these articles were published between 2020–2021. In total, 74% of studies employed PRISMA or other methods to report their SLR. A quality assessment was not conducted in 117 of the 220 included studies (Table 1). A review of the ICD codes revealed that neoplasms (Chapter II) were the most frequently studied clinical areas, followed by diseases of the nervous system (Chapter VI) (Table 2). As far as the data sources used are concerned, imaging was used most frequently with clinical notes and lab tests following as the second frequently used (Table 2).

Considerable variations were observed across the included publications in terms of ML accuracy, specificity, and sensitivity. ICD-10 Chapters III and XVIII reported the lowest results, while some ML algorithms for ICD-10 II and VI reported 100% accuracy for all three parameters (Appendix 1). In total, 231 of 10,963studies (7 of 220 SLR) provided information about the accuracy, specificity, and sensitivity of all included studies. A total of 3,164 studies (51 SLRs) did not report any results across the three dimensions (Table 3).

Four thousand nine hundred ninety-two of the 10,963studies (103 of 220 SLRs) conducted internal validation procedures. The most common approach was the k-fold cross-validation (1,325 studies), followed by leave-one-out cross-validation (205 cases) (Appendix 2).

Regarding external validation, a comparison of ML with a human comparator was mentioned in 90 of the 10,963studies (Table 4) (227-313). In total, 50 cases provided evidence of comparable performance, 33 (four) publications confirmed the superiority (inferiority) of ML over clinicians and three did not indicate any results. The median number of clinical experts included in the validation was six (range, 1–511).

The methodological approach to the missing data was discussed in 144 studies, with the most common being imputation (Table 5).

In total, over 10,000 ML algorithms were used for the included SLRs (Table 6). The most common modelling approach was neural networks (2,454 studies), followed by SVM and RF/decision trees (1,578 and 1,522 studies, respectively).

Discussion

To the best of our knowledge, this is the first attempt at systematically studying the integration of ML algorithms in healthcare.

The key finding was the low reporting quality of publications dedicated to the development and adaptation of ML algorithms in clinical practice. There was a significant share of studies without data on accuracy (44%), sensitivity (72%), and specificity (75%), as well as internal (65%) and external (99%) validations. Additionally, 10, 819 of the 10, 963 studies (98%) did not report a methodological approach for missing data.

The majority of publications across the 220 SLRs aimed at ML solutions towards clinical prediction and categorisation challenges as well as disease prognosis in oncology and neurology using mainly imaging data. Natural language processing (NLP) was employed mainly in the SLR reporting applications in Chapter XXI, in which clinicians' notes were more likely to be used as data sources. This is consistent with a recent study that revealed that 189 (85%) of 222 FDA-approved medical devices were intended for use by healthcare professionals, while 33 (15%) were intended for use by patients (314).

However, the most frequently published type of ML is the artificial neural network (ANN). ANNs can detect complex nonlinear relationships and interactions between the dependent and independent variables (universal approximators). Deep learning (DL) methods are primarily used in oncological or respiratory disease studies. The increasing use of DL has been observed during the COVID-19 pandemic. A systematic literature review of 34 studies indicated that ML can enhance the sensitivity and specificity of radiographic images compared with radiologists' diagnoses.

Our review indicated that apart from neural networks, SVM is the most frequently used after deep neural networks. The highly similar performance of SVM, particularly in terms of classification accuracy, makes them rank among the most popular ML classifiers. In addition to deep networks and SVM, RF is most often used. It is an ensemble bagging technique whereby numerous decision trees are combined to obtain the results. This process combines bootstrapping and aggregation. The key advantage of this approach is that it can be used for classification and regression problems; hence, this is likely one of the reasons it is used in cardiology. Although RF can provide higher diagnostic accuracy and reduce variance without increasing bias, the operating time might be extensive in clinical situations. However, boosted methods known to improve the performance of the corresponding methods have not been extensively encountered.

This study has several limitations. First, our review was limited to literature reviews; consequently, certain information might have been misunderstood if it had not been presented in a given SLR. There may have been some over-counting of the number of ML algorithms identified. We did not have sufficient details to understand whether any of the publications used the same data source. Second, we did not review studies that were missed in any SLR; hence, we could have a biased picture of the utilisation of ML in healthcare. Third, we restricted the review to studies published in English, thus potentially introducing a selection bias towards particular countries. Finally, publication bias cannot be excluded, as reports of unsatisfactory/unsuccessful ML applications are rarely encountered; thus, the actual performance of ML could be overestimated. Several techniques may be reported differently and may be missed or incorrectly categorised. For example, principal component analysis (PCA) was also reported in the included SLRs, despite not being strictly an ML algorithm but a dimensionality reduction technique.

Finally, the main focus in evaluating ML performance was accuracy/sensitivity/specificity, which is relevant in classification problems such as disease diagnosis; however, prediction models such as length of stay (LoS) can be evaluated through other parameters such as residual mean square error (RMSE) or the coefficient of determination ($R^{2)}$.

Despite these limitations, this review provides important insights into the current state of AI integration in the healthcare sector. This indicates that over 10,000 ML algorithms have been developed for healthcare systems. This is not surprising, considering that AI is becoming a major driver of innovation in healthcare. For example, the number of patents granted solely for digital communication or medical technologies will almost double that for drugs by 2022 in Europe (315). A rough comparison indicates that fewer than 60 drugs and over 100 ML/AI-enabled medical devices have been approved annually by the FDA since 2019 (up to 523 until the end of January 2023) (316).

There is an evident gap between the development and utilisation of ML algorithms in healthcare, prompting us to elicit recommendations for both developers and payers.

Recommendations for ML developers in healthcare

With respect to accuracy, specificity, and sensitivity, the results of our review appear promising at first glance. It must be considered an impressive picture, with 12 therapeutic areas (out of 22 ICD chapters) having access to ML algorithms with an accuracy of 100%. In addition, five ICD chapters had scores above 88% (Appendix 1). However, to understand the clinical usability of AI, a review of such absolute numbers may not provide a full picture. The adoption of ML algorithms in clinical settings requires further validation. The lack of testing of the predictive power on separate datasets may overestimate ML performance in practical situations. Validation and cross-validation lead to more accurate estimates of the performance of the ML model on an unseen dataset. This was the most common approach, but it was found in only 15% of the included 10,462 studies (Appendix 1). Cross-validation divides the sample into k subsets, with k subsets used as the test set/validation set and k-1 subsets for training. The model was trained on the training data and predictions were made using the model on the testing data. The

sensitivity and/or specificity were averaged by testing multiple times on k-fold data subsets. As most of the data were used for fitting, the k-fold approach significantly reduced the bias and variance, as most of the data were also used in the validation set. That is, it reduces the risk of undertraining when a large amount of noise is introduced into the training data and, consequently, bias. It helps prevent overfitting, which occurs when the model attempts to learn each detail and noise of the data, leading to poor model performance on test sets (317).

The adoption of an appropriate methodology for validation is the bare minimum to approach the concept of effectiveness introduced for pharmaceuticals. An adequate amount of data is also important for ensuring the best ML accuracy, specificity, and sensitivity. The cross-validation performed better with larger datasets. This is vital, particularly when one considers the importance of ML for diagnosis which initiates the sequence of subsequent actions. Hence, it is up to the correct prediction that we can allow the healthcare system to be effective and efficient. This helps to effectively optimise treatment pathways for previously diagnosed patients. The availability of data enables healthcare professionals to use predictive modelling techniques for prevention and prophylaxis actions more than ever.

Generally, the larger the dataset, the greater the statistical power and chances of better prediction. A negative relationship between sample size and classification accuracy has already been reported (317, 318), and it is important to note that as many as 83 out of 220 systematic literature reviews did not provide information regarding the size of the datasets used for ML algorithm development. Simultaneously, the majority of the included SLRs reported a large variance between the smallest and largest sample sizes, despite having similar clinical objectives (Table 1). However, it is not only the size of the training dataset that has significant importance, but also the variability of the available data. Insufficient diversity in training datasets can lead to inadequate generalisation of the model outputs to different patient populations. Training models on multi-institutional datasets can be the most effective in combating model deterioration, and directly combating existing biases in training data can also mitigate their impact. Some studies have indicated that other sources of potential variety driven by medical device manufacturer software are adopted, in which AI models trained on cardiac magnetic resonance imaging (MRI) scans provide different accuracy results from different scanners (319), and more than two-fold differences were found in the error rate between two different optical coherence tomography (OCT) scans (320). The limited diversity in the data used for ML is a problem, and a scoping review of publications related to AI that appeared in PubMed in 2019 revealed that over half of the datasets used for clinical AI originated from either the US or China (321). In addition, the U.S. and China contributed over 40% of the publications (321). Barriers to accessing data lead to the overutilisation of available datasets. For instance, there are only four major databases in ophthalmology, ESSIDOR, DRIVE, EyePACS, and E-ophtha, with unknown publicly available datasets for ophthalmological images in 172 countries that constitute roughly 45% of the global population (322).

Cross-validation also helps to ensure that the best ML model is used for a given decision problem. When studies published comparative data, the results indicate that results differ for different ML methods used on the same datasets. For instance, across 12 studies using deep neural networks for ECG analysis to

detect structural cardiac pathologies, the predictive accuracy of the neural network DL models was superior to that of expert interpretations by board-certified cardiologists. The same was found in the comparison of computer-aided detection (CAD) systems with 53 general endoscopists for detecting early neoplasia in patients with Barrett's oesophagus (BE). The CAD achieved higher accuracy than any of the clinicians, regardless of the level of endoscopic expertise (323); in both cases, details regarding the choice of the clinical group were missing.

It should be noted that less than 1% of the included studies reported external validation. This is a significant gap in the evidence. This has two specific consequences for its implementation in clinical practice. For true external validation, a tuned algorithm must be applied to a new set of data from different sources. The ultimate objective was to ensure the generalisability of the results with the adoption of ML across various care compositions. As Bang and colleagues mentioned in their systematic literature review: "CAD algorithms demonstrated high accuracy for the automatic endoscopic diagnosis of oesophageal cancer and neoplasms. The limitation of a lack of performance in external validation and clinical applications should be overcome" (20).

Our findings are similar to those of another review of DL studies that focused on the comparison of ML against human comparators covering the period from 2010 to June 2022. Only ten RCTs (including eight ongoing RCTs) and 81 non-randomised clinical trials compared diagnostic algorithm performance against clinicians (324). In another systematic literature review of 82 publications, only 14 studies compared the diagnostic performance of DL models based on medical imaging with that of healthcare professionals (325).

Recommendations for regulators and payers

Will improvements in both internal and external validations make ML algorithms directly eligible for registration and refundable? While the former is likely more about internal validity, as its primary objective addresses the risk–benefit ratio, the latter may be more about external validity, as its primary objective is to address the value for money. Therefore, the next question is how regulators and public payers should balance the requirements with respect to the evidence of the usability of ML algorithms against the need to ensure safety and treatment effectiveness. First, given the existence of strict regulations for both market regulation and pricing and reimbursement for pharmaceuticals and medical devices, it is necessary to enquire whether similar hurdles of evidence generation should also be introduced for ML algorithms. To address this issue, we introduce two random facts: only approximately 12% of drugs entering clinical trials are ultimately approved by the FDA (326), and the average time to reimbursement for innovative treatments in Europe is 511 days (327). Hence, some claim that overregulation may harm innovation. However, the development of the majority of AI-driven innovations may be relatively short compared to other time-consuming research and development technologies, and there is potential for greater disruption in the healthcare sector by ML algorithms than what we have witnessed thus far. Therefore, the types of regulations that should be developed to support the adoption of ML algorithms remain unclear. Overall, there is a need to establish a matrix of criteria to assess the ability of AI solutions to be integrated into healthcare systems. There are already several recommendations in this respect, such

as a scoping review of 72 guidelines that, among others, identified quality criteria regarding the development, evaluation, and implementation of ML in healthcare (328). Other experts have suggested grouping ML algorithms into one of the following categories: assistive, augmentative, or autonomous (329).

Still, there is a need for decision-makers (regulators and public payers) to form a common unified approach towards the development of a common set of standards for the assessment of AI-driven health technologies, as ML is seldom jurisdiction-specific. The maximum accuracy varied from 27% (ICD XVIII) to 100% (ICD II) across the included studies. Therefore, the question is whether the same rules should be applied, irrespective of the area under consideration. This may require the involvement of clinical experts and a clear understanding of the unmet medical needs in each disease field. Therefore, our recommendations first focus on interoperability in the journey towards unified P&R regulations for ML algorithms. The underlying rationale is to ensure the accessibility of data such as electronic medical records (EMRs) to AI developers. Thus far, there have been limited efforts related to the availability of real-world data (RWD) for validation as eluded earlier. In the era of digital transformation, we should move further and ensure the integration of EMRs with unstructured data. Additionally, healthcare decision-makers must prepare data repositories to facilitate external validation and invest in local data analytics capabilities to facilitate internal validation. Such efforts should be welcomed by developers, as expressed by many experts (330). The overarching objective is to ensure that ML algorithms have complete access to health-related data irrespective of geographical, demographic, or institutional composition. Without an appropriate understanding of the health problems in question, ML algorithms can only be utilised for the populations and medical conditions for which they were trained, failing to provide any value for populations or concomitant medical conditions that were omitted or underrepresented in the training set owing to racial, ethnic, or simple misrepresentation. Such activities will inevitably bring an additional burden on both payers and developers; however, AI is as good as the data it possesses, as demonstrated in this study.

Conclusions

There is still unrealised potential for AI in healthcare. Despite the growing number of published ML algorithms, there is limited evidence of their impact on clinical practice.

More evidence related to external and internal validation can drive the change towards a greater, more robust, and safer adoption of AI. Consequently, it may allow payers, clinicians, and patients to increase their trust in ML algorithms. The key is ensuring that AI development is examined through the lens of the health problems in question. Unmet medical needs are heterogeneously shaped by patients and influenced by the care setting, baseline characteristics, and cultural differences. Thus, there is a need to prepare a landing field for ML algorithms for healthcare applications. However, we are not there yet. Hence, by moving forward, AI will only face more challenges. Currently, we are in a different era. Let us be ready with the right data at the appropriate time.

Declarations

Ethical Approval

Not applicable

Competing interests

None

Authors' contributions

K.K. and S.P. wrote the main manuscript text and B.Y. M.H and K.K. prepared all tables and figure. All authors reviewed the manuscript

Funding

Not applicable

Availability of data and materials

Not applicable

References

1. OECD. Health at a Glance 2021: OECD Indicators: Digital health., 2021.
2. Zhang A, Xing L, Zou J, et al. Shifting machine learning for healthcare from development to deployment and from models to data. Nat Biomed Eng. 2022; 6: 1330-45.
3. Loftus TJ, Tighe PJ, Ozrazgat-Baslanti T, et al. Ideal algorithms in healthcare: Explainable, dynamic, precise, autonomous, fair, and reproducible. PLOS Digit Health. 2022; 1.
4. Padula WV, Kreif N, Vanness DJ, et al. Machine Learning Methods in Health Economics and Outcomes Research-The PALISADE Checklist: A Good Practices Report of an ISPOR Task Force. Value Health. 2022; 25: 1063-80.
5. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009; 339: b2535.
6. Abu Bakar AR, Lai KW, Hamzaid NA. The emergence of machine learning in auditory neural impairment: A systematic review. Neurosci Lett. 2021; 765: 136250.
7. Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: A systematic review. Comput Struct Biotechnol J. 2021; 19: 2833-50.
8. Adeoye J, Tan JY, Choi SW, et al. Prediction models applying machine learning to oral cavity cancer outcomes: A systematic review. INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS. 2021; 154.

9. Ahsan MM, Siddique Z. Machine learning-based heart disease diagnosis: A systematic literature review. Artificial Intelligence in Medicine. 2022; 128.
10. Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: Current status and future challenges-A systematic review. ARTIFICIAL INTELLIGENCE IN MEDICINE. 2021; 120.
11. Al Hinai G, Jammoul S, Vajihi Z, et al. Deep learning analysis of resting electrocardiograms for the detection of myocardial dysfunction, hypertrophy, and ischaemia: a systematic review. Eur Heart J Digit Health. 2021; 2: 416-23.
12. Alabi RO, Youssef O, Pirinen M, et al. Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. Artificial Intelligence in Medicine. 2021; 115.
13. Albahri AS, Hamid RA, Alwan J, et al. Role of biological Data Mining and Machine Learning Techniques in Detecting and Diagnosing the Novel Coronavirus (COVID-19): A Systematic Review. Journal of Medical Systems. 2020; 44.
14. Alballa N, Al-Turaiki I. Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. Informatics in Medicine Unlocked. 2021; 24.
15. Alharbi ET, Nadeem F, Cherif A. Predictive models for personalized asthma attacks based on patient's biosignals and environmental factors: a systematic review. BMC MEDICAL INFORMATICS AND DECISION MAKING. 2021; 21.
16. Alhasan AS. Clinical Applications of Artificial Intelligence, Machine Learning, and Deep Learning in the Imaging of Gliomas: A Systematic Review. CUREUS. 2021; 13.
17. Alsolai H, Qureshi S, Iqbal SMZ, et al. A Systematic Review of Literature on Automated Sleep Scoring. IEEE Access. 2022; 10: 79419-43.
18. Anteby R, Klang E, Horesh N, et al. Deep learning for noninvasive liver fibrosis classification: A systematic review. Liver Int. 2021; 41: 2269-78.
19. Bang CS, Lee JJ, Baik GH. Computer-Aided Diagnosis of Gastrointestinal Ulcer and Hemorrhage Using Wireless Capsule Endoscopy: Systematic Review and Diagnostic Test Accuracy Meta-analysis. J Med Internet Res. 2021; 23: e33267.
20. Bang CS, Lee JJ, Baik GH. Computer-aided diagnosis of esophageal cancer and neoplasms in endoscopic images: a systematic review and meta-analysis of diagnostic test accuracy. Gastrointest Endosc. 2021; 93: 1006-15.e13.
21. Barrett L, Hu J, Howell P. Systematic Review of Machine Learning Approaches for Detecting Developmental Stuttering. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2022; 30: 1160-72.
22. Bazoukis G, Stavrakis S, Zhou J, et al. Machine learning versus conventional clinical methods in guiding management of heart failure patients—a systematic review. Heart Failure Reviews. 2021; 26: 23-34.
23. Bedrikovetski S, Dudi-Venkata NN, Kroon HM, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. BMC Cancer. 2021; 21: 1058.

24. Bedrikovetski S, Dudi-Venkata NN, Maicas G, et al. Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: A systematic review and meta-analysis. Artif Intell Med. 2021; 113: 102022.
25. Benoit J, Onyeaka H, Keshavan M, et al. Systematic Review of Digital Phenotyping and Machine Learning in Psychosis Spectrum Illnesses. Harv Rev Psychiatry. 2020; 28: 296-304.
26. Bernert RA, Hilberg AM, Melia R, et al. Artificial Intelligence and Suicide Prevention: A Systematic Review of Machine Learning Investigations. Int J Environ Res Public Health. 2020; 17.
27. Bertl M, Metsallik J, Ross P. A systematic literature review of AI-based digital decision support systems for post-traumatic stress disorder. Front Psychiatry. 2022; 13.
28. Binvignat M, Pedoia V, Butte AJ, et al. Use of machine learning in osteoarthritis research: A systematic literature review. RMD Open. 2022; 8.
29. Boonstra A, Laven M. Influence of artificial intelligence on the work design of emergency department clinicians a systematic literature review. BMC Health Serv Res. 2022; 22.
30. Boyd C, Brown G, Kleinig T, et al. Machine Learning Quantitation of Cardiovascular and Cerebrovascular Disease: A Systematic Review of Clinical Applications. Diagnostics (Basel). 2021; 11.
31. Bracher-Smith M, Crawford K, Escott-Price V. Machine learning for genetic prediction of psychiatric disorders: a systematic review. Mol Psychiatry. 2021; 26: 70-79.
32. Buchlak QD, Esmaili N, Leveque JC, et al. Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review. Neurosurg Rev. 2020; 43: 1235-53.
33. Buisson M, Navel V, Labbe A, et al. Deep learning versus ophthalmologists for screening for glaucoma on fundus examination: A systematic review and meta-analysis. CLINICAL AND EXPERIMENTAL OPHTHALMOLOGY. 2021; 49: 1027-38.
34. Cabitza F, Locoro A, Banfi G. Machine Learning in Orthopedics: A Literature Review. Front Bioeng Biotechnol. 2018; 6: 75.
35. Castaldo R, Cavaliere C, Soricelli A, et al. Radiomic and Genomic Machine Learning Method Performance for Prostate Cancer Diagnosis: Systematic Literature Review. J Med Internet Res. 2021; 23: e22394.
36. Cavus N, Lawan AA, Ibrahim Z, et al. A Systematic Literature Review on the Application of Machine-Learning Models in Behavioral Assessment of Autism Spectrum Disorder. J Pers Med. 2021; 11.
37. Celtikci E. A systematic review on machine learning in neurosurgery: The future of decision-making in patient care. Turkish Neurosurgery. 2018; 28: 167-73.
38. Chandra G, Irisha KD, Vica VI, et al. Systematic Literature Review on Application of Artificial Intelligence in Cancer Detection Using Image Processing. 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS), 2022.

39. Chee ML, Ong MEH, Siddiqui FJ, et al. Artificial Intelligence Applications for COVID-19 in Intensive Care and Emergency Settings: A Systematic Review. Int J Environ Res Public Health. 2021; 18.
40. Chiesa-Estomba CM, Graña M, Medela A, et al. Machine Learning Algorithms as a Computer-Assisted Decision Tool for Oral Cancer Prognosis and Management Decisions: A Systematic Review. ORL J Otorhinolaryngol Relat Spec. 2022: 1-11.
41. Cho SJ, Sunwoo L, Baik SH, et al. Brain metastasis detection using machine learning: a systematic review and meta-analysis. Neuro Oncol. 2021; 23: 214-25.
42. Choudhury A, Asan O. Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review. JMIR Med Inform. 2020; 8: e18599.
43. Choudhury A, Renjilian E, Asan O. Use of machine learning in geriatric clinical care for chronic diseases: a systematic literature review. JAMIA Open. 2020; 3: 459-71.
44. da Silva Neto SR, Tabosa Oliveira T, Teixeira IV, et al. Machine learning and deep learning techniques to support clinical diagnosis of arboviral diseases: A systematic review. PLoS Negl Trop Dis. 2022; 16: e0010061.
45. Dallora AL, Eivazzadeh S, Mendes E, et al. Machine learning and microsimulation techniques on the prognosis of dementia: A systematic literature review. PLoS One. 2017; 12: e0179804.
46. Dallora AL, Anderberg P, Kvist O, et al. Bone age assessment with various machine learning techniques: A systematic literature review and meta-analysis. PLoS One. 2019; 14: e0220242.
47. Daniel, Cenggoro TW, Pardamean B. A systematic literature review of machine learning application in COVID-19 medical image classification. Procedia computer science. 2023; 216: 749-56.
48. D'Antoni F, Russo F, Ambrosio L, et al. Artificial Intelligence and Computer Vision in Low Back Pain: A Systematic Review. Int J Environ Res Public Health. 2021; 18.
49. Das PK, D. V A, Meher S, et al. A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia. IEEE Access. 2022; 10: 81741-63.
50. Das T, Kaur H, Gour P, et al. Intersection of network medicine and machine learning towards investigating the key biomarkers and pathways underlying amyotrophic lateral sclerosis: a systematic review. Briefings in bioinformatics. 2022; 23.
51. de Bardeci M, Ip CT, Olbrich S. Deep learning applied to electroencephalogram data in mental disorders: A systematic review. Biol Psychol. 2021; 162: 108117.
52. Decharatanachart P, Chaiteerakij R, Tiyarattanachai T, et al. Application of artificial intelligence in chronic liver diseases: a systematic review and meta-analysis. BMC Gastroenterol. 2021; 21: 10.
53. Decharatanachart P, Chaiteerakij R, Tiyarattanachai T, et al. Application of artificial intelligence in non-alcoholic fatty liver disease and liver fibrosis: a systematic review and meta-analysis. Therap Adv Gastroenterol. 2021; 14: 17562848211062807.
54. DelSole EM, Keck WL, Patel AA. The State of Machine Learning in Spine Surgery: A Systematic Review. Clin Spine Surg. 2021.

55. Dogan O, Tiwari S, Jabbar MA, et al. A systematic review on AI/ML approaches against COVID-19 outbreak. Complex Intell Systems. 2021: 1-24.
56. Dudchenko A, Kopanitsa G. Decision Support Systems in Cardiology: A Systematic Review. Studies in Health Technology & Informatics. 2017; 237: 209-14.
57. Ebrahimi A, Wiil UK, Schmidt T, et al. Predicting the Risk of Alcohol Use Disorder Using Machine Learning: A Systematic Literature Review. IEEE Access. 2021; 9: 151697-712.
58. Ebrahimighahnavieh MA, Luo S, Chiong R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. Comput Methods Programs Biomed. 2020; 187: 105242.
59. El-Daw S, El-Tantawy A, Aly T, et al. Role of machine learning in management of degenerative spondylolisthesis: A systematic review. Current Orthopaedic Practice. 2021; 32: 302-08.
60. Falconer N, Abdel-Hafez A, Scott IA, et al. Systematic review of machine learning models for personalised dosing of heparin. BRITISH JOURNAL OF CLINICAL PHARMACOLOGY. 2021; 87: 4124-39.
61. Farook TH, Jamayet NB, Abdullah JY, et al. Machine Learning and Intelligent Diagnostics in Dental and Orofacial Pain Management: A Systematic Review. Pain Research and Management. 2021; 2021.
62. Fernandes F, Barbalho I, Barros D, et al. Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review. BioMedical Engineering Online. 2021; 20.
63. Fregoso-Aparicio L, Noguez J, Montesinos L, et al. Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. Diabetol Metab Syndr. 2021; 13: 148.
64. Frondelius T, Atkova I, Miettunen J, et al. Diagnostic and prognostic prediction models in ventilator-associated pneumonia: Systematic review and meta-analysis of prediction modelling studies. JOURNAL OF CRITICAL CARE. 2022; 67: 44-56.
65. Fusco R, Grassi R, Granata V, et al. Artificial intelligence and covid-19 using chest ct scan and chest x-ray images: Machine learning and deep learning approaches for diagnosis and treatment. Journal of Personalized Medicine. 2021; 11.
66. Garrow CR, Kowalewski KF, Li L, et al. Machine Learning for Surgical Phase Recognition: A Systematic Review. Ann Surg. 2021; 273: 684-93.
67. Ghaderzadeh M, Asadi F. Deep Learning in the Detection and Diagnosis of COVID-19 Using Radiology Modalities: A Systematic Review. J Healthc Eng. 2021; 2021: 6677314.
68. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimers Res Ther. 2021; 13: 162.
69. Gutiérrez-Tobal GC, Álvarez D, Kheirandish-Gozal L, et al. Reliability of machine learning to diagnose pediatric obstructive sleep apnea: Systematic review and meta-analysis. Pediatr Pulmonol. 2021.

70. Haggenmüller S, Maron RC, Hekler A, et al. Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts. Eur J Cancer. 2021; 156: 202-16.
71. Hameed BMZ, Shah M, Naik N, et al. The Ascent of Artificial Intelligence in Endourology: a Systematic Review Over the Last 2 Decades. Curr Urol Rep. 2021; 22: 53.
72. Hasan N, Bao YKK. Understanding current states of machine learning approaches in medical informatics: a systematic literature review. HEALTH AND TECHNOLOGY. 2021; 11: 471-82.
73. Hassan N, Slight R, Weiand D, et al. Preventing sepsis; how can artificial intelligence inform the clinical decision-making process? A systematic review. Int J Med Inform. 2021; 150: 104457.
74. Henn J, Buness A, Schmid M, et al. Machine learning to guide clinical decision-making in abdominal surgery-a systematic literature review. Langenbecks Arch Surg. 2021.
75. Hickman SE, Woitek R, Le EPV, et al. Machine Learning for Workflow Applications in Screening Mammography: Systematic Review and Meta-Analysis. Radiology. 2022; 302: 88-104.
76. Hinterwimmer F, Lazic I, Suren C, et al. Machine learning in knee arthroplasty: specific data are key-a systematic review. Knee Surg Sports Traumatol Arthrosc. 2022.
77. Hoekstra O, Hurst W, Tummers J. Healthcare related event prediction from textual data with machine learning: A Systematic Literature Review. Healthc Anal. 2022; 2.
78. Hoodbhoy Z, Masroor Jeelani S, Aziz A, et al. Machine Learning for Child and Adolescent Health: A Systematic Review. Pediatrics. 2021; 147.
79. Hosni M, Abnane I, Idri A, et al. Reviewing ensemble classification methods in breast cancer. Computer Methods and Programs in Biomedicine. 2019; 177: 89-112.
80. Hoyos W, Aguilar J, Toro M. Dengue models based on machine learning techniques: A systematic literature review. Artif Intell Med. 2021; 119: 102157.
81. Huang J, Shlobin NA, DeCuypere M, et al. Deep Learning for Outcome Prediction in Neurosurgery: A Systematic Review of Design, Reporting, and Reproducibility. Neurosurgery. 2022; 90: 16-38.
82. Huang S, Dang J, Sheckter CC, et al. A systematic review of machine learning and automation in burn wound evaluation: A promising but developing frontier. Burns. 2021; 47: 1691-704.
83. Huang Z, Aarab G, Ravesloot MJL, et al. Prediction of the obstruction sites in the upper airway in sleep-disordered breathing based on snoring sound parameters: a systematic review. Sleep Med. 2021; 88: 116-33.
84. Ibrahim B, Suppiah S, Ibrahim N, et al. Diagnostic power of resting-state fMRI for detection of network connectivity in Alzheimer's disease and mild cognitive impairment: A systematic review. Hum Brain Mapp. 2021; 42: 2941-68.
85. Infante T, Cavaliere C, Punzo B, et al. Radiogenomics and Artificial Intelligence Approaches Applied to Cardiac Computed Tomography Angiography and Cardiac Magnetic Resonance for Precision Medicine in Coronary Heart Disease: A Systematic Review. Circ Cardiovasc Imaging. 2021; 14: 1133-46.

86. Irgang L, Barth H, Holmén M. Data-Driven Technologies as Enablers for Value Creation in the Prevention of Surgical Site Infections: a Systematic Review. J Healthc Informatics Res. 2023; 7: 1-41.
87. Islam MN, Mustafina SN, Mahmud T, et al. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. BMC Pregnancy Childbirth. 2022; 22.
88. Jiang K, Jiang X, Pan J, et al. Current Evidence and Future Perspective of Accuracy of Artificial Intelligence Application for Early Gastric Cancer Diagnosis With Endoscopy: A Systematic and Meta-Analysis. Front Med (Lausanne). 2021; 8: 629080.
89. Jiang MY, Ma YX, Guo SY, et al. Using Machine Learning Technologies in Pressure Injury Management: Systematic Review. JMIR MEDICAL INFORMATICS. 2021; 9.
90. Jones OT, Calanzani N, Saji S, et al. Artificial Intelligence Techniques That May Be Applied to Primary Care Data to Facilitate Earlier Diagnosis of Cancer: Systematic Review. Journal of Medical Internet Research. 2021; 23: N.PAG-N.PAG.
91. Kalhori SRN, Tanhapour M, Gholamzadeh M. Enhanced childhood diseases treatment using computational models: Systematic review of intelligent experiments heading to precision medicine. JOURNAL OF BIOMEDICAL INFORMATICS. 2021; 115.
92. Kareemi H, Vaillancourt C, Rosenberg H, et al. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021; 28: 184-96.
93. Karwath A, Bunting KV, Gill SK, et al. Redefining β-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet. 2021; 398: 1427-35.
94. Kassem MA, Hosny KM, Damaševičius R, et al. Machine Learning and Deep Learning Methods for Skin Lesion Classification and Diagnosis: A Systematic Review. Diagnostics (Basel). 2021; 11.
95. Kausch SL, Moorman JR, Lake DE, et al. Physiological machine learning models for prediction of sepsis in hospitalized adults: An integrative review. Intensive and Critical Care Nursing. 2021; 65.
96. Kawamoto A, Takenaka K, Okamoto R, et al. Systematic review of artificial intelligence-based image diagnosis for inflammatory bowel disease. Digestive endoscopy : official journal of the Japan Gastroenterological Endoscopy Society. 2022; 34: 1311-19.
97. Kedra J, Radstake T, Pandit A, et al. Current status of use of big data and artificial intelligence in RMDs: a systematic literature review informing EULAR recommendations. RMD Open. 2019; 5: e001004.
98. Kennedy EE, Bowles KH, Aryal S. Systematic review of prediction models for postacute care destination decision-making. J Am Med Informatics Assoc. 2022; 29: 176-86.
99. Khanagar SB, Naik S, Al Kheraif AA, et al. Application and Performance of Artificial Intelligence Technology in Oral Cancer Diagnosis and Prediction of Prognosis: A Systematic Review. Diagnostics (Basel). 2021; 11.
100. Kim HR, Sung M, Park JA, et al. Analyzing adverse drug reaction using statistical and machine learning methods: A systematic review. Medicine. 2022; 101: E29387.

101. Kim SS. Recent trends of artificial intelligence and machine learning for insomnia research. Chronobiology in Medicine. 2021; 3: 16-19.
102. Kodama S, Fujihara K, Shiozaki H, et al. Ability of current machine learning algorithms to predict and detect hypoglycemia in patients with diabetes mellitus: meta-analysis. JMIR Diabetes. 2021; 6.
103. Komolafe TE, Cao Y, Nguchu BA, et al. Diagnostic Test Accuracy of Deep Learning Detection of COVID-19: A Systematic Review and Meta-Analysis. Acad Radiol. 2021; 28: 1507-23.
104. Kourou K, Exarchos KP, Papaloukas C, et al. Applied machine learning in cancer research: A systematic review for patient diagnosis, classification and prognosis. Comput Struct Biotechnol J. 2021; 19: 5546-55.
105. Kozikowski M, Suarez-Ibarrola R, Osiecki R, et al. Role of Radiomics in the Prediction of Muscle-invasive Bladder Cancer: A Systematic Review and Meta-analysis. Eur Urol Focus. 2021.
106. Kumar S, Oh I, Schindler S, et al. Machine learning for modeling the progression of Alzheimer disease dementia using clinical data: a systematic literature review. JAMIA Open. 2021; 4: ooab052.
107. Kumar Y, Gupta S, Singla R, et al. A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. ARCHIVES OF COMPUTATIONAL METHODS IN ENGINEERING.
108. Kuntz S, Krieghoff-Henning E, Kather JN, et al. Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review. Eur J Cancer. 2021; 155: 200-15.
109. La Greca Saint-Esteven A, Vuong D, Tschanz F, et al. Systematic Review on the Association of Radiomics with Tumor Biological Endpoints. Cancers (Basel). 2021; 13.
110. Langarizadeh M, Sayadi M. Machine Learning Techniques for Diagnosis of Lower Gastrointestinal Cancer: A Systematic Review. IRANIAN RED CRESCENT MEDICAL JOURNAL. 2021; 23.
111. Le Glaz A, Haralambous Y, Kim-Dufor DH, et al. Machine Learning and Natural Language Processing in Mental Health: Systematic Review. J Med Internet Res. 2021; 23: e15708.
112. Lecointre L, Dana J, Lodi M, et al. Artificial intelligence-based radiomics models in endometrial cancer: A systematic review. Eur J Surg Oncol. 2021; 47: 2734-41.
113. Lequertier V, Wang T, Fondrevelle J, et al. Hospital Length of Stay Prediction Methods: A Systematic Review. Med Care. 2021; 59: 929-38.
114. Li JX, Zhou ZJ, Dong JY, et al. Predicting breast cancer 5-year survival using machine learning: A systematic review. PLOS ONE. 2021; 16.
115. Li MD, Ahmed SR, Choy E, et al. Artificial intelligence applied to musculoskeletal oncology: a systematic review. SKELETAL RADIOLOGY. 2022; 51: 245-56.
116. Li Y, Wang X, Zhang J, et al. Applications of artificial intelligence (AI) in researches on non-alcoholic fatty liver disease(NAFLD) : A systematic review. Rev Endocr Metab Disord. 2021.
117. Librenza-Garcia D, Kotzian BJ, Yang J, et al. The impact of machine learning techniques in the study of bipolar disorder: A systematic review. NEUROSCIENCE AND BIOBEHAVIORAL REVIEWS. 2017; 80: 538-54.

118. Lima CLD, da Silva ACG, Moreno GMM, et al. Temporal and Spatiotemporal Arboviruses Forecasting by Machine Learning: A Systematic Review. Front Public Health. 2022; 10.
119. Locquet M, Diep AN, Beaudart C, et al. A systematic review of prediction models to diagnose COVID-19 in adults admitted to healthcare centers. Arch Public Health. 2021; 79: 105.
120. Lopez CD, Gazgalis A, Boddapati V, et al. Artificial Learning and Machine Learning Decision Guidance Applications in Total Hip and Knee Arthroplasty: A Systematic Review. Arthroplast Today. 2021; 11: 103-12.
121. Lubelski D, Hersh A, Azad TD, et al. Prediction Models in Degenerative Spine Surgery: A Systematic Review. Global Spine J. 2021; 11: 79s-88s.
122. Maile H, Li JPO, Gore D, et al. Machine Learning Algorithms to Detect Subclinical Keratoconus: Systematic Review. JMIR MEDICAL INFORMATICS. 2021; 9.
123. Mangold C, Zoretic S, Thallapureddy K, et al. Machine Learning Models for Predicting Neonatal Mortality: A Systematic Review. Neonatology. 2021; 118: 394-405.
124. Mari T, Henderson J, Maden M, et al. Systematic Review of the Effectiveness of Machine Learning Algorithms for Classifying Pain Intensity, Phenotype or Treatment Outcomes Using Electroencephalogram Data. J Pain. 2021.
125. Matsangidou M, Liampas A, Pittara M, et al. Machine Learning in Pain Medicine: An Up-To-Date Systematic Review. Pain Ther. 2021; 10: 1067-84.
126. Mawdsley E, Reynolds B, Cullen B. A systematic review of the effectiveness of machine learning for predicting psychosocial outcomes in acquired brain injury: Which algorithms are used and why? J Neuropsychol. 2021; 15: 319-39.
127. Medic G, Kosaner Kließ M, Atallah L, et al. Evidence-based Clinical Decision Support Systems for the prediction and detection of three disease states in critical care: A systematic literature review. F1000Res. 2019; 8: 1728.
128. Mei J, Desrosiers C, Frasnelli J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. Front Aging Neurosci. 2021; 13: 633752.
129. Mellia JA, Basta MN, Toyoda Y, et al. Natural Language Processing in Surgery A Systematic Review and Meta-analysis. ANNALS OF SURGERY. 2021; 273: 900-08.
130. Miltiadous A, Tzimourta KD, Giannakeas N, et al. Machine Learning Algorithms for Epilepsy Detection Based on Published EEG Databases: A Systematic Review. IEEE Access. 2023; 11: 564-94.
131. Minissi ME, Chicchi Giglioli IA, Mantovani F, et al. Assessment of the Autism Spectrum Disorder Based on Machine Learning and Social Visual Attention: A Systematic Review. J Autism Dev Disord. 2021.
132. Miranda L, Paul R, Pütz B, et al. Systematic Review of Functional MRI Applications for Psychiatric Disease Subtyping. Front Psychiatry. 2021; 12: 665536.
133. Mirzania D, Thompson AC, Muir KW. Applications of deep learning in detection of glaucoma: A systematic review. Eur J Ophthalmol. 2021; 31: 1618-42.

134. Moezzi M, Shirbandi K, Shahvandi HK, et al. The diagnostic accuracy of Artificial Intelligence-Assisted CT imaging in COVID-19 disease: A systematic review and meta-analysis. Inform Med Unlocked. 2021; 24: 100591.
135. Moglia A, Georgiou K, Georgiou E, et al. A systematic review on artificial intelligence in robot-assisted surgery. Int J Surg. 2021; 95: 106151.
136. Mohan BP, Khan SR, Kassab LL, et al. High pooled performance of convolutional neural networks in computer-aided diagnosis of GI ulcers and/or hemorrhage on wireless capsule endoscopy images: a systematic review and meta-analysis. Gastrointest Endosc. 2021; 93: 356-64.e4.
137. Mondal MRH, Bharati S, Podder P. Diagnosis of COVID-19 Using Machine Learning and Deep Learning: A Review. CURRENT MEDICAL IMAGING. 2021; 17: 1403-18.
138. Montazeri M, ZahediNasab R, Farahani A, et al. Machine Learning Models for Image-Based Diagnosis and Prognosis of COVID-19: Systematic Review. JMIR Med Inform. 2021; 9: e25181.
139. Moor M, Rieck B, Horn M, et al. Early Prediction of Sepsis in the ICU Using Machine Learning: A Systematic Review. Front Med (Lausanne). 2021; 8: 607952.
140. Moshawrab M, Adda M, Bouzouane A, et al. Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review. Sensors (Basel, Switzerland). 2023; 23.
141. Moura FSE, Amin K, Ekwobi C. Artificial intelligence in the management and treatment of burns: a systematic review. BURNS & TRAUMA. 2021; 9.
142. Mpanya D, Celik T, Klug E, et al. Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. Int J Cardiol Heart Vasc. 2021; 34: 100773.
143. Mughal H, Javed AR, Rizwan M, et al. Parkinson's Disease Management via Wearable Sensors: A Systematic Review. IEEE Access. 2022; 10: 35219-37.
144. Musa N, Gital AYu, Aljojo N, et al. A systematic review and Meta-data analysis on the applications of Deep Learning in Electrocardiogram. Journal of ambient intelligence and humanized computing. 2022: 1-74.
145. Musulin J, Baressi Šegota S, Štifanić D, et al. Application of Artificial Intelligence-Based Regression Methods in the Problem of COVID-19 Spread Prediction: A Systematic Review. Int J Environ Res Public Health. 2021; 18.
146. Nadarajah R, Alsaeed E, Hurdus B, et al. Prediction of incident atrial fibrillation in community-based electronic health records: a systematic review with meta-analysis. Heart. 2021.
147. Naemi A, Schmidt T, Mansourvar M, et al. Machine learning techniques for mortality prediction in emergency departments: a systematic review. BMJ OPEN. 2021; 11.
148. Nafea MS, Ismail ZH. Supervised Machine Learning and Deep Learning Techniques for Epileptic Seizure Recognition Using EEG Signals-A Systematic Literature Review. Bioengineering (Basel, Switzerland). 2022; 9.
149. Nasser M, Yusof UK. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. Diagnostics (Basel, Switzerland). 2023; 13.

150. Nazarian S, Glover B, Ashrafian H, et al. Diagnostic Accuracy of Artificial Intelligence and Computer-Aided Diagnosis for the Detection and Characterization of Colorectal Polyps: Systematic Review and Meta-analysis. J Med Internet Res. 2021; 23: e27370.
151. Nguyen AV, Blears EE, Ross E, et al. Machine learning applications for the differentiation of primary central nervous system lymphoma from glioblastoma on imaging: a systematic review and meta-analysis. Neurosurg Focus. 2018; 45: E5.
152. Ogink PT, Groot OQ, Bindels BJJ, et al. The use of machine learning prediction models in spinal surgical outcome: An overview of current development and external validation studies. Seminars in Spine Surgery. 2021; 33.
153. Ogink PT, Groot OQ, Karhade AV, et al. Wide range of applications for machine-learning prediction models in orthopedic surgical outcome: a systematic review. Acta Orthop. 2021; 92: 526-31.
154. Ortíz-Barrios MA, Coba-Blanco DM, Alfaro-Saíz JJ, et al. Process Improvement Approaches for Increasing the Response of Emergency Departments against the COVID-19 Pandemic: A Systematic Review. Int J Environ Res Public Health. 2021; 18.
155. Ossai CI, Wickramasinghe N. Intelligent decision support with machine learning for efficient management of mechanical ventilation in the intensive care unit – A critical overview. International Journal of Medical Informatics. 2021; 150.
156. Paganelli AI, Mondéjar AG, da Silva AC, et al. Real-time data analysis in health monitoring systems: A comprehensive systematic literature review. Journal of Biomedical Informatics. 2022; 127.
157. Pahwa B, Bali O, Goyal S, et al. Applications of Machine Learning in Pediatric Hydrocephalus: A Systematic Review. Neurology India. 2021; 69: S568-S77.
158. Patil S, Habib Awan K, Arakeri G, et al. Machine learning and its potential applications to the genomic study of head and neck cancer—A systematic review. Journal of Oral Pathology and Medicine. 2019; 48: 773-79.
159. Peralta M, Jannin P, Baxter JSH. Machine learning in deep brain stimulation: A systematic review*. ARTIFICIAL INTELLIGENCE IN MEDICINE. 2021; 122.
160. Persad E, Jost K, Honoré A, et al. Neonatal sepsis prediction through clinical decision support algorithms: A systematic review. Acta Paediatr. 2021; 110: 3201-26.
161. Popa SL, Ismaiel A, Cristina P, et al. Non-Alcoholic Fatty Liver Disease: Implementing Complete Automated Diagnosis and Staging. A Systematic Review. Diagnostics (Basel). 2021; 11.
162. Prasoppokakorn T, Tiyarattanachai T, Chaiteerakij R, et al. Application of artificial intelligence for diagnosis of pancreatic ductal adenocarcinoma by EUS: A systematic review and meta-analysis. Endosc Ultrasound. 2021.
163. Quaak M, van de Mortel L, Thomas RM, et al. Deep learning applications for the classification of psychiatric disorders using neuroimaging data: Systematic review and meta-analysis. Neuroimage Clin. 2021; 30: 102584.
164. Quartuccio N, Marrale M, Laudicella R, et al. The role of PET radiomic features in prostate cancer: a systematic review. CLINICAL AND TRANSLATIONAL IMAGING. 2021; 9: 579-88.

165. Ramesh S, Chokkara S, Shen T, et al. Applications of Artificial Intelligence in Pediatric Oncology: A Systematic Review. JCO Clin Cancer Inform. 2021; 5: 1208-19.
166. Ramos-Lima LF, Waikamp V, Antonelli-Salgado T, et al. The use of machine learning techniques in trauma-related disorders: a systematic review. J Psychiatr Res. 2020; 121: 159-72.
167. Ravegnini G, Ferioli M, Morganti AG, et al. Radiomics and Artificial Intelligence in Uterine Sarcomas: A Systematic Review. J Pers Med. 2021; 11.
168. Ren M, Yi PH. Artificial intelligence in orthopedic implant model classification: a systematic review. Skeletal Radiol. 2022; 51: 407-16.
169. Rice P, Pugh M, Geraghty R, et al. Machine Learning Models for Predicting Stone-Free Status after Shockwave Lithotripsy: A Systematic Review and Meta-Analysis. Urology. 2021; 156: 16-22.
170. Rowe TW, Katzourou IK, Stevenson-Hoare JO, et al. Machine learning for the life-time risk prediction of Alzheimer's disease: a systematic review. Brain Commun. 2021; 3: fcab246.
171. Safaei M, Sundararajan EA, Driss M, et al. A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. Comput Biol Med. 2021; 136: 104754.
172. Sajjadian M, Lam RW, Milev R, et al. Machine learning in the prediction of depression treatment outcomes: a systematic review and meta-analysis. Psychol Med. 2021; 51: 2742-51.
173. Salas-Zárate R, Alor-Hernández G, Salas-Zárate MDP, et al. Detecting Depression Signs on Social Media: A Systematic Literature Review. Healthcare (Basel, Switzerland). 2022; 10.
174. Salem H, Soria D, Lund JN, et al. A systematic review of the applications of Expert Systems (ES) and machine learning (ML) in clinical urology. BMC Med Inform Decis Mak. 2021; 21: 223.
175. Sanmarchi F, Fanconi C, Golinelli D, et al. Predict, diagnose, and treat chronic kidney disease with machine learning: a systematic literature review. J Nephrol. 2023.
176. Sanmarchi F, Fanconi C, Golinelli D, et al. Correction to: Predict, diagnose, and treat chronic kidney disease with machine learning: a systematic literature review (Journal of Nephrology, (2023), 10.1007/s40620-023-01573-4). J Nephrol. 2023.
177. Saputro SA, Pattanaprateep O, Pattanateepapon A, et al. Prognostic models of diabetic microvascular complications: a systematic review and meta-analysis. Syst Rev. 2021; 10: 288.
178. Sardar SK, Kumar N, Lee SC. A Systematic Literature Review on Machine Learning Algorithms for Human Status Detection. IEEE Access. 2022; 10: 74366-82.
179. Scardoni A, Balzarini F, Signorelli C, et al. Artificial intelligence-based tools to control healthcare associated infections: A systematic review of the literature. J Infect Public Health. 2020; 13: 1061-77.
180. Segato A, Marzullo A, Calimeri F, et al. Artificial intelligence for brain diseases: A systematic review. APL Bioeng. 2020; 4: 041503.
181. Senanayake S, White N, Graves N, et al. Machine learning in predicting graft failure following kidney transplantation: A systematic review of published predictive models. INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS. 2019; 130.

182. Senders JT, Arnaout O, Karhade AV, et al. Natural and Artificial Intelligence in Neurosurgery: A Systematic Review. Neurosurgery. 2018; 83: 181-92.
183. Shi Z, Zhang Z, Liu Z, et al. Methodological quality of machine learning-based quantitative imaging analysis studies in esophageal cancer: a systematic review of clinical outcome prediction after concurrent chemoradiotherapy. Eur J Nucl Med Mol Imaging. 2021.
184. Shillan D, Sterne JAC, Champneys A, et al. Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. Crit Care. 2019; 23: 284.
185. Shin S, Austin PC, Ross HJ, et al. Machine learning vs. conventional statistical models for predicting heart failure readmission and mortality. ESC Heart Failure. 2021; 8: 106-15.
186. Shlobin NA, Baig AA, Waqas M, et al. Artificial Intelligence for Large-Vessel Occlusion Stroke: A Systematic Review. World Neurosurg. 2021.
187. Siddiqui S, Arifeen M, Hopgood A, et al. Deep Learning Models for the Diagnosis and Screening of COVID-19: A Systematic Review. SN computer science. 2022; 3: 397.
188. Smets J, Shevroja E, Hügle T, et al. Machine Learning Solutions for Osteoporosis—A Review. Journal of Bone and Mineral Research. 2021; 36: 833-51.
189. Soffer S, Klang E, Shimon O, et al. Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis. Sci Rep. 2021; 11: 15814.
190. Song DY, Topriceanu CC, Ilie-Ablachim DC, et al. Machine learning with neuroimaging data to identify autism spectrum disorder: a systematic review and meta-analysis. NEURORADIOLOGY. 2021; 63: 2057-72.
191. Song X, Liu X, Liu F, et al. Comparison of machine learning and logistic regression models in predicting acute kidney injury: A systematic review and meta-analysis. Int J Med Inform. 2021; 151: 104484.
192. Sorin V, Barash Y, Konen E, et al. Deep Learning for Natural Language Processing in Radiology-Fundamentals and a Systematic Review. JOURNAL OF THE AMERICAN COLLEGE OF RADIOLOGY. 2020; 17: 639-48.
193. Srivani M, Murugappan A, Mala T. Cognitive computing technological trends and future research directions in healthcare — A systematic literature review. Artificial Intelligence in Medicine. 2023; 138.
194. Stephens ME, O'Neal CM, Westrup AM, et al. Utility of machine learning algorithms in degenerative cervical and lumbar spine disease: a systematic review. Neurosurgical Review. 2021.
195. Stewart J, Lu J, Goudie A, et al. Applications of machine learning to undifferentiated chest pain in the emergency department: A systematic review. PLoS One. 2021; 16: e0252612.
196. Stokes K, Castaldo R, Federici C, et al. The use of artificial intelligence systems in diagnosis of pneumonia via signs and symptoms: A systematic review. Biomed Signal Process Control. 2022; 72.
197. Subramanian H, Dey R, Brim WR, et al. Trends in Development of Novel Machine Learning Methods for the Identification of Gliomas in Datasets That Include Non-Glioma Images: A Systematic Review.

Front Oncol. 2021; 11: 788819.

198. Syeda HB, Syed M, Sexton KW, et al. Role of Machine Learning Techniques to Tackle the COVID-19 Crisis: Systematic Review. JMIR Med Inform. 2021; 9: e23811.
199. Syer T, Mehta P, Antonelli M, et al. Artificial Intelligence Compared to Radiologists for the Initial Diagnosis of Prostate Cancer on Magnetic Resonance Imaging: A Systematic Review and Recommendations for Future Studies. Cancers (Basel). 2021; 13.
200. Tabatabaei M, Razaei A, Sarrami AH, et al. Current Status and Quality of Machine Learning-Based Radiomics Studies for Glioma Grading: A Systematic Review. Oncology. 2021; 99: 433-43.
201. Tan KR, Seng JJB, Kwan YH, et al. Evaluation of Machine Learning Methods Developed for Prediction of Diabetes Complications: A Systematic Review. J Diabetes Sci Technol. 2021: 19322968211056917.
202. Teo YH, Lim ICZY, Tseng FS, et al. Predicting Clinical Outcomes in Acute Ischemic Stroke Patients Undergoing Endovascular Thrombectomy with Machine Learning: A Systematic Review and Meta-analysis. Clinical Neuroradiology. 2021; 31: 1121-30.
203. Tewarie IA, Senders JT, Kremer S, et al. Survival prediction of glioblastoma patients-are we there yet? A systematic review of prognostic modeling for glioblastoma and its clinical potential. Neurosurg Rev. 2021; 44: 2047-57.
204. Triantafyllidis A, Polychronidou E, Alexiadis A, et al. Computerized decision support and machine learning applications for the prevention and treatment of childhood obesity: A systematic review of the literature. ARTIFICIAL INTELLIGENCE IN MEDICINE. 2020; 104.
205. Ugga L, Perillo T, Cuocolo R, et al. Meningioma MRI radiomics and machine learning: systematic review, quality score assessment, and meta-analysis. Neuroradiology. 2021; 63: 1293-304.
206. van Kempen EJ, Post M, Mannil M, et al. Accuracy of Machine Learning Algorithms for the Classification of Molecular Features of Gliomas on MRI: A Systematic Literature Review and Meta-Analysis. Cancers (Basel). 2021; 13.
207. van Kempen EJ, Post M, Mannil M, et al. Performance of machine learning algorithms for glioma segmentation of brain MRI: a systematic literature review and meta-analysis. Eur Radiol. 2021; 31: 9638-53.
208. Volpe S, Pepa M, Zaffaroni M, et al. Machine Learning for Head and Neck Cancer: A Safe Bet?-A Clinically Oriented Systematic Review for the Radiation Oncologist. Front Oncol. 2021; 11: 772663.
209. Wang W, Kiik M, Peek N, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. PLoS One. 2020; 15: e0234722.
210. Wesselius FJ, van Schie MS, De Groot NMS, et al. Digital biomarkers and algorithms for detection of atrial fibrillation using surface electrocardiograms: A systematic review. Comput Biol Med. 2021; 133: 104404.
211. Wongkoblap A, Vadillo MA, Curcin V. Researching Mental Health Disorders in the Era of Social Media: Systematic Review. J Med Internet Res. 2017; 19: e228.

212. Wu JH, Liu TYA, Hsu WT, et al. Performance and limitation of machine learning algorithms for diabetic retinopathy screening: Meta-analysis. Journal of Medical Internet Research. 2021; 23.
213. Xu L, He B, Zhang Y, et al. Prognostic models for amyotrophic lateral sclerosis: a systematic review. J Neurol. 2021; 268: 3361-70.
214. Yan MY, Gustad LT, Nytrø Ø. Sepsis prediction, early detection, and identification using clinical text for machine learning: a systematic review. J Am Med Inform Assoc. 2021.
215. Yeo M, Tahayori B, Kok HK, et al. Review of deep learning algorithms for the automatic detection of intracranial hemorrhages on computed tomography head imaging. J Neurointerv Surg. 2021; 13: 369-78.
216. Yin J, Ngiam KY, Teo HH. Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review. J Med Internet Res. 2021; 23: e25759.
217. Yu K, Syed MN, Bernardis E, et al. Machine Learning Applications in the Evaluation and Management of Psoriasis: A Systematic Review. J Psoriasis Psoriatic Arthritis. 2020; 5: 147-59.
218. Zakhem GA, Fakhoury JW, Motosko CC, et al. Characterizing the role of dermatologists in developing artificial intelligence for assessment of skin cancer. Journal of the American Academy of Dermatology. 2021; 85: 1544-56.
219. Zaunseder E, Haupt S, Mütze U, et al. Opportunities and challenges in machine learning-based newborn screening-A systematic literature review. JIMD Reports. 2022; 63: 250-61.
220. Zeiser FA, da Costa CA, Roehe AV, et al. Breast cancer intelligent analysis of histopathological data: A systematic review. APPLIED SOFT COMPUTING. 2021; 113.
221. Zhang L, Sun JQ, Jiang BB, et al. Development of artificial intelligence in epicardial and pericoronary adipose tissue imaging: a systematic review. EUROPEAN JOURNAL OF HYBRID IMAGING. 2021; 5.
222. Zhao Y, Wood EP, Mirin N, et al. Social Determinants in Machine Learning Cardiovascular Disease Prediction Models: A Systematic Review. Am J Prev Med. 2021; 61: 596-605.
223. Zheng QH, Yang L, Zeng B, et al. Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis. ECLINICALMEDICINE. 2021; 31.
224. Zheng Y, Dickson VV, Blecker S, et al. Identifying Patients With Hypoglycemia Using Natural Language Processing: Systematic Literature Review. JMIR diabetes. 2022; 7: e34681.
225. Zhou Y, Ge YT, Shi XL, et al. Machine learning predictive models for acute pancreatitis: A systematic review. Int J Med Inform. 2022; 157: 104641.
226. Zhu T, Li K, Herrero P, et al. Deep Learning for Diabetes: A Systematic Review. IEEE J Biomed Health Inform. 2021; 25: 2744-57.
227. Ariji Y, Fukuda M, Kise Y, et al. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence. Oral Surg Oral Med Oral Pathol Oral Radiol. 2019; 127: 458-63.

228. Ariji Y, Sugita Y, Nagao T, et al. CT evaluation of extranodal extension of cervical lymph node metastases in patients with oral squamous cell carcinoma using deep learning classification. Oral Radiol. 2020; 36: 148-55.
229. Kwon JM, Jeon KH, Kim HM, et al. Comparing the performance of artificial intelligence and conventional diagnosis criteria for detecting left ventricular hypertrophy using electrocardiography. Europace. 2020; 22: 412-19.
230. Nakajima K, Kudo T, Nakata T, et al. Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study. Eur J Nucl Med Mol Imaging. 2017; 44: 2280-89.
231. Al-Aswad LA, Kapoor R, Chu CK, et al. Evaluation of a Deep Learning System For Identifying Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. Journal of Glaucoma. 2019; 28: 1029 - 34.
232. Liu S, Graham SL, Schulz A, et al. A Deep Learning-Based Algorithm Identifies Glaucomatous Discs Using Monoscopic Fundus Photographs. Ophthalmology Glaucoma. 2018; 1: 15-22.
233. Phene S, Dunn RC, Hammel N, et al. Deep Learning and Glaucoma Specialists: The Relative Importance of Optic Disc Features to Predict Glaucoma Referral in Fundus Photographs. Ophthalmology. 2019; 126: 1627-39.
234. Shibata N, Tanito M, Mitsuhashi K, et al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. Sci Rep. 2018; 8: 14665.
235. Seo E, Jaccard N, Trikha S, et al. Automated Evaluation of Optic Disc Images for Manifest Glaucoma Detection Using a Deep-Learning, Neural Network-Based Algorithm. Investigative Ophthalmology & Visual Science. 2018; 59: 2080-80.
236. Jammal AA, Thompson AC, Mariottoni EB, et al. Human Versus Machine: Comparing a Deep Learning Algorithm to Human Gradings for Detecting Glaucoma on Fundus Photographs. Am J Ophthalmol. 2020; 211: 123-31.
237. Brinker TJ, Hekler A, Enk AH, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur J Cancer. 2019; 113: 47-54.
238. Brinker TJ, Hekler A, Enk AH, et al. Deep neural networks are superior to dermatologists in melanoma image classification. Eur J Cancer. 2019; 119: 11-17.
239. Yu C, Yang S, Kim W, et al. Acral melanoma detection using a convolutional neural network for dermoscopy images. PLoS One. 2018; 13: e0193321.
240. Marchetti MA, Codella NCF, Dusza SW, et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. J Am Acad Dermatol. 2018; 78: 270-77.e1.
241. Marchetti MA, Liopyris K, Dusza SW, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: Results of the International Skin Imaging Collaboration 2017. J Am Acad Dermatol. 2020; 82: 622-27.

242. Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol. 2018; 29: 1836-42.
243. Haenssle HA, Fink C, Toberer F, et al. Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. Annals of Oncology. 2020; 31: 137-43.
244. Haenssle HA, Winkler JK, Fink C, et al. Skin lesions of face and scalp - Classification by a market-approved convolutional neural network in comparison with 64 dermatologists. Eur J Cancer. 2021; 144: 192-99.
245. Tschandl P, Codella N, Akay BN, et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. Lancet Oncol. 2019; 20: 938-47.
246. Maron RC, Weichenthal M, Utikal JS, et al. Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. Eur J Cancer. 2019; 119: 57-65.
247. Tschandl P, Rosendahl C, Akay BN, et al. Expert-Level Diagnosis of Nonpigmented Skin Cancer by Combined Convolutional Neural Networks. JAMA Dermatol. 2019; 155: 58-65.
248. Fujisawa Y, Otomo Y, Ogata Y, et al. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. Br J Dermatol. 2019; 180: 373-81.
249. Jinnai S, Yamazaki N, Hirano Y, et al. The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning. Biomolecules. 2020; 10.
250. Han SS, Park I, Eun Chang S, et al. Augmented Intelligence Dermatology: Deep Neural Networks Empower Medical Professionals in Diagnosing Skin Cancer and Predicting Treatment Options for 134 Skin Disorders. J Invest Dermatol. 2020; 140: 1753-61.
251. Han SS, Kim MS, Lim W, et al. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. J Invest Dermatol. 2018; 138: 1529-38.
252. Brinker TJ, Hekler A, Enk AH, et al. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer. 2019; 111: 148-54.
253. Han SS, Moon IJ, Kim SH, et al. Assessment of deep neural networks for the diagnosis of benign and malignant skin neoplasms in comparison with dermatologists: A retrospective validation study. PLoS Med. 2020; 17: e1003381.
254. Hekler A, Utikal JS, Enk AH, et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. Eur J Cancer. 2019; 118: 91-96.
255. Brinker TJ, Schmitt M, Krieghoff-Henning EI, et al. Diagnostic performance of artificial intelligence for histologic melanoma recognition compared to 18 international expert pathologists. J Am Acad

Dermatol. 2022; 86: 640-42.

256. Brennan M, Puri S, Ozrazgat-Baslanti T, et al. Comparing clinical judgment with the MySurgeryRisk algorithm for preoperative risk assessment: A pilot usability study. Surgery. 2019; 165: 1035-45.
257. Kambakamba P, Mannil M, Herrera PE, et al. The potential of machine learning to predict postoperative pancreatic fistula based on preoperative, non-contrast-enhanced CT: A proof-of-principle study. Surgery. 2020; 167: 448-54.
258. Yala A, Schuster T, Miles R, et al. A Deep Learning Model to Triage Screening Mammograms: A Simulation Study. Radiology. 2019; 293: 38-46.
259. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020; 577: 89-94.
260. Balta C, Rodríguez-Ruiz A, Mieskes C, et al. Going from double to single reading for screening exams labeled as likely normal by AI: what is the impact? , Other Conferences, 2020.
261. Dembrower K, Wåhlin E, Liu Y, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. Lancet Digit Health. 2020; 2: e468-e74.
262. Kyono T, Gilbert FJ, van der Schaar M. Improving Workflow Efficiency for Mammography Using Machine Learning. J Am Coll Radiol. 2020; 17: 56-63.
263. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. Eur Radiol. 2019; 29: 4825-32.
264. Geras KJ, Wolfson S, Kim SG, et al. High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks. ArXiv. 2017; abs/1703.07047.
265. Lotter W, Diab AR, Haslam B, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. Nat Med. 2021; 27: 244-49.
266. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology. 2019; 290: 305-14.
267. Schaffter T, Buist DSM, Lee CI, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. JAMA Netw Open. 2020; 3: e200265.
268. Salim M, Wåhlin E, Dembrower K, et al. External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. JAMA Oncol. 2020; 6: 1581-88.
269. Kim HE, Kim HH, Han BK, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. Lancet Digit Health. 2020; 2: e138-e48.
270. Cowley JB. The use of knowledge discovery databases in the identification of patients with colorectal cancer. 2012.

271. Wei JW, Suriawinata AA, Vaickus LJ, et al. Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathologic Slides. JAMA Netw Open. 2020; 3: e203398.
272. Song Z, Yu C, Zou S, et al. Automatic deep learning-based colorectal adenoma detection system and its similarities with pathologists. BMJ Open. 2020; 10: e036423.
273. Bychkov D, Linder N, Turkki R, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep. 2018; 8: 3395.
274. Geessink OGF, Baidoshvili A, Klaase JM, et al. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. Cell Oncol (Dordr). 2019; 42: 331-41.
275. Kather JN, Krisam J, Charoentong P, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS Med. 2019; 16: e1002730.
276. Zhao K, Li Z, Yao S, et al. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer. EBioMedicine. 2020; 61: 103054.
277. Suh HB, Choi YS, Bae S, et al. Primary central nervous system lymphoma and atypical glioblastoma: Differentiation using radiomics approach. Eur Radiol. 2018; 28: 3832-39.
278. Kang D, Park JE, Kim YH, et al. Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma: development and multicenter external validation. Neuro Oncol. 2018; 20: 1251-61.
279. Alcaide-Leon P, Dufort P, Geraldo AF, et al. Differentiation of Enhancing Glioma and Primary Central Nervous System Lymphoma by Texture-Based Machine Learning. AJNR Am J Neuroradiol. 2017; 38: 1145-50.
280. Yamashita K, Yoshiura T, Arimura H, et al. Performance evaluation of radiologists with artificial neural network for differential diagnosis of intra-axial cerebral tumors on MR images. AJNR Am J Neuroradiol. 2008; 29: 1153-8.
281. Rehm GB, Han J, Kuhn BT, et al. Creation of a Robust and Generalizable Machine Learning Classifier for Patient Ventilator Asynchrony. Methods Inf Med. 2018; 57: 208-19.
282. Bakkes T, Montree RJH, Mischi M, et al. A machine learning method for automatic detection and classification of patient-ventilator asynchrony. Annu Int Conf IEEE Eng Med Biol Soc. 2020; 2020: 150-53.
283. Mulqueeny Q, Redmond SJ, Tassaux D, et al. Automated detection of asynchrony in patient-ventilator interaction. Annu Int Conf IEEE Eng Med Biol Soc. 2009; 2009: 5324-7.
284. Aghdam MA, Sharifi A, Pedram MM. Diagnosis of Autism Spectrum Disorders in Young Children Based on Resting-State Functional Magnetic Resonance Imaging Data Using Convolutional Neural Networks. J Digit Imaging. 2019; 32: 899-918.
285. Petrucci K, Petrucci P, Canfield K, et al. Evaluation of UNIS: Urological Nursing Information Systems. Proc Annu Symp Comput Appl Med Care. 1991: 43-7.
286. Gorman R. Expert system for management of urinary incontinence in women. Proc Annu Symp Comput Appl Med Care. 1995: 527-31.

287. Chang PL, Li YC, Wang TM, et al. Evaluation of a decision-support system for preoperative staging of prostate cancer. Med Decis Making. 1999; 19: 419-27.
288. Koutsojannis C, Hatzilygeroudis I. FESMI: A Fuzzy Expert System for Diagnosis and Treatment of Male Impotence. International Conference on Knowledge-Based Intelligent Information & Engineering Systems, 2004.
289. Koutsojannis C, Nabil E, Tsimara M, et al. Using Machine Learning Techniques to Improve the Behaviour of a Medical Decision Support System for Prostate Diseases. 2009.
290. Altunay S, Telatar Z, Eroğul O, et al. A new approach to urinary system dynamics problems: Evaluation and classification of uroflowmeter signals using artificial neural networks. Expert Syst Appl. 2009; 36: 4891-95.
291. Koutsojannis C, Lithari C, Hatzilygeroudis I. Managing urinary incontinence through hand-held real-time decision support aid. Comput Methods Programs Biomed. 2012; 107: 84-9.
292. Hassanien AE, Alqaheri H, El-Dahshan E-S. Prostate boundary detection in ultrasound images using biologically-inspired spiking neural network. Applied Soft Computing. 2011; 11: 2035-41.
293. Torshizi AD, Zarandi MH, Torshizi GD, et al. A hybrid fuzzy-ontology based intelligent system to determine level of severity and treatment recommendation for Benign Prostatic Hyperplasia. Comput Methods Programs Biomed. 2014; 113: 301-13.
294. Xiao D, Zhang G, Liu Y, et al. 3D detection and extraction of bladder tumors via MR virtual cystoscopy. Int J Comput Assist Radiol Surg. 2016; 11: 89-97.
295. Hurst RE, Bonner RB, Ashenayi K, et al. Neural net-based identification of cells expressing the p300 tumor-related antigen using fluorescence image analysis. Cytometry. 1997; 27: 36-42.
296. Hao AT, Wu LP, Kumar A, et al. Nursing process decision support system for urology ward. Int J Med Inform. 2013; 82: 604-12.
297. Lopes MH, Ortega NR, Silveira PS, et al. Fuzzy cognitive map in differential diagnosis of alterations in urinary elimination: a nursing approach. Int J Med Inform. 2013; 82: 201-8.
298. Koutsojannis C, Tsimara M, Nabil E. HIROFILOS: a medical expert system for prostate diseases. 2008.
299. Qiu W, Kuang H, Teleg E, et al. Machine Learning for Detecting Early Infarction in Acute Stroke with Non-Contrast-enhanced CT. Radiology. 2020; 294: 638-44.
300. Tomita N, Cheung YY, Hassanpour S. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. Comput Biol Med. 2018; 98: 8-15.
301. Murata K, Endo K, Aihara T, et al. Artificial intelligence for the detection of vertebral fractures on plain spinal radiography. Sci Rep. 2020; 10: 20031.
302. Cheng CT, Ho TY, Lee TY, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. Eur Radiol. 2019; 29: 5469-77.
303. Yu JS, Yu SM, Erdal BS, et al. Detection and localisation of hip fractures on anteroposterior radiographs with artificial intelligence: proof of concept. Clin Radiol. 2020; 75: 237.e1-37.e9.

304. Yamada Y, Maki S, Kishida S, et al. Automated classification of hip fractures using deep convolutional neural networks with orthopedic surgeon-level accuracy: ensemble decision-making with antero-posterior and lateral radiographs. Acta Orthop. 2020; 91: 699-704.
305. Jiménez-Sánchez A, Kazi A, Albarqouni S, et al. Precise proximal femur fracture classification for interactive training and surgical planning. Int J Comput Assist Radiol Surg. 2020; 15: 847-57.
306. Adams M, Chen W, Holcdorf D, et al. Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures. J Med Imaging Radiat Oncol. 2019; 63: 27-32.
307. Mawatari T, Hayashida Y, Katsuragawa S, et al. The effect of deep convolutional neural networks on radiologists' performance in the detection of hip fractures on digital pelvic radiographs. Eur J Radiol. 2020; 130: 109188.
308. Urakawa T, Tanaka Y, Goto S, et al. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiol. 2019; 48: 239-44.
309. Chung SW, Han SS, Lee JW, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop. 2018; 89: 468-73.
310. Olczak J, Fahlberg N, Maki A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs. Acta Orthop. 2017; 88: 581-86.
311. Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A. 2018; 115: 11591-96.
312. Qazi AA, Pekar V, Kim J, et al. Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach. Med Phys. 2011; 38: 6160-70.
313. Misra-Hebert AD, Milinovich A, Zajichek A, et al. Natural Language Processing Improves Detection of Nonsevere Hypoglycemia in Medical Records Versus Coding Alone in Patients With Type 2 Diabetes but Does Not Improve Prediction of Severe Hypoglycemia Events: An Analysis Using the Electronic Medical Record in a Large Health System. Diabetes Care. 2020; 43: 1937-40.
314. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. The Lancet Digital Health. 2021; 3: e195-e203.
315. Office EP.
316. FDA. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. 2022.
317. Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage. 2018; 180: 68-77.
318. Vabalas A, Gowen E, Poliakoff E, et al. Machine learning algorithm validation with a limited sample size. PLOS ONE. 2019; 14: e0224365.
319. Campello VM, Gkontra P, Izquierdo C, et al. Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge. IEEE Trans Med Imaging. 2021; 40: 3543-54.
320. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018; 24: 1342-50.

321. Celi LA, Cellini J, Charpignon M-L, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. PLOS Digital Health. 2022; 1: e0000022.
322. Khan SM, Liu X, Nath S, et al. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. Lancet Digit Health. 2021; 3: e51-e66.
323. de Groof AJ, Struyvenberg MR, van der Putten J, et al. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. Gastroenterology. 2020; 158: 915-29.e4.
324. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. Bmj. 2020; 368: m689.
325. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital Health. 2019; 1: e271-e97.
326. Congress. CBONaftUS. Research and Development in the Pharmaceutical Industry. 2023.
327. Associations. EFooPIa. The root cause of unavailability and delay to innovative medicines: Reducing the time before patients have access to innovative medicines., 2022.
328. de Hond AAH, Leeuwenberg AM, Hooft L, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. NPJ Digit Med. 2022; 5: 2.
329. Parikh RB, Helmchen LA. Paying for artificial intelligence in medicine. NPJ Digit Med. 2022; 5: 63.
330. Abràmoff MD, Roehrenbeck C, Trujillo S, et al. A reimbursement framework for artificial intelligence in healthcare. NPJ Digit Med. 2022; 5: 72.
331. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021; 372: n71.
332. Maniah, Soewito B, Lumban Gaol F, et al. A systematic literature Review: Risk analysis in cloud migration. Journal of King Saud University - Computer and Information Sciences. 2022; 34: 3111-20.
333. Kitchenham B, Pearl Brereton O, Budgen D, et al. Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology. 2009; 51: 7-15.
334. Kitchenham B, Charters S. Guidelines for Performing Systematic Literature Reviews in Software Engineering, Technical Report EBSE 2007-001. 2007.
335. Brereton P, Kitchenham BA, Budgen D, et al. Lessons from applying the systematic literature review process within the software engineering domain. Journal of Systems and Software. 2007; 80: 571-83.
336. Luo W, Phung D, Tran T, et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. J Med Internet Res. 2016; 18: e323.
337. Zeng X, Zhang Y, Kwong JS, et al. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. J Evid Based Med. 2015; 8: 2-10.

338. Kable AK, Pich J, Maslin-Prothero SE. A structured approach to documenting a search strategy for publication: a 12 step guideline for authors. Nurse Educ Today. 2012; 32: 878-86.
339. Kim HR, Sung M, Park JA, et al. Analyzing adverse drug reaction using statistical and machine learning methods: A systematic review. Medicine. 2022; 101: e29387.
340. Paganelli AI, Mondéjar AG, da Silva AC, et al. Real-time data analysis in health monitoring systems: A comprehensive systematic literature review. J Biomed Informatics. 2022; 127: 104009.
341. Ni Q, Sun ZY, Qi L, et al. A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images. European Radiology. 2020; 30: 6517-27.

Tables

Tables 1 to 6 are available in the Supplementary Files section

Figures

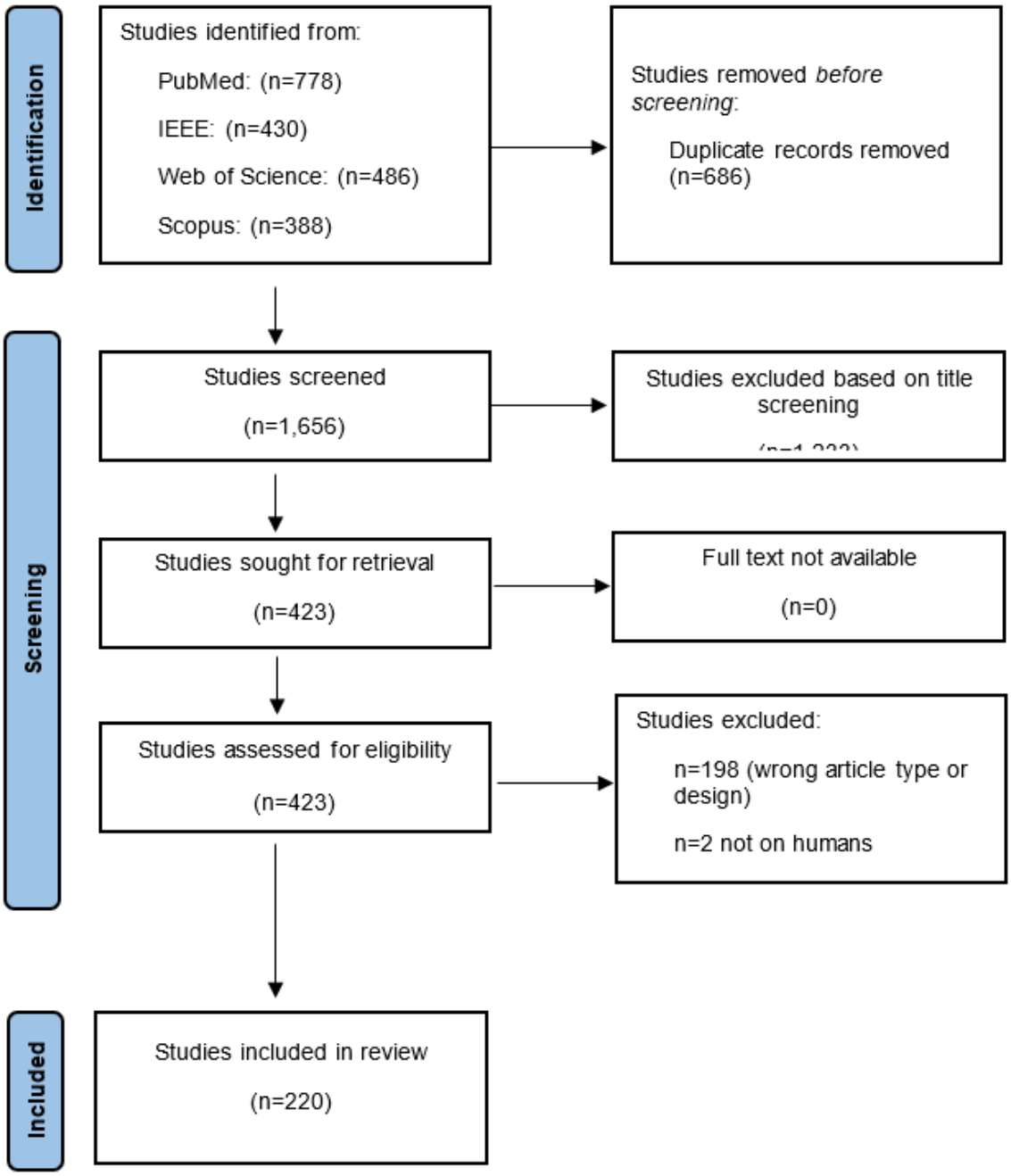

Figure 1

PRISMA flowchart for selection of publications (331).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- mlAppendix19.04.docx
- mlTables19.04.docx