

Review

Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review

Jiamin Yin^{1*}, BA; Kee Yuan Ngiam^{2*}, MBBS; Hock Hai Teo^{1*}, PhD

¹Department of Information Systems and Analytics, School of Computing, National University of Singapore, Singapore, Singapore

²Department of Surgery, National University Hospital, Singapore, Singapore

* all authors contributed equally

Corresponding Author:

Hock Hai Teo, PhD

Department of Information Systems and Analytics

School of Computing

National University of Singapore

13 Computing Drive, NUS

Singapore, 117417

Singapore

Phone: 65 65162979

Email: teohh@comp.nus.edu.sg

Abstract

Background: Artificial intelligence (AI) applications are growing at an unprecedented pace in health care, including disease diagnosis, triage or screening, risk analysis, surgical operations, and so forth. Despite a great deal of research in the development and validation of health care AI, only few applications have been actually implemented at the frontlines of clinical practice.

Objective: The objective of this study was to systematically review AI applications that have been implemented in real-life clinical practice.

Methods: We conducted a literature search in PubMed, Embase, Cochrane Central, and CINAHL to identify relevant articles published between January 2010 and May 2020. We also hand searched premier computer science journals and conferences as well as registered clinical trials. Studies were included if they reported AI applications that had been implemented in real-world clinical settings.

Results: We identified 51 relevant studies that reported the implementation and evaluation of AI applications in clinical practice, of which 13 adopted a randomized controlled trial design and eight adopted an experimental design. The AI applications targeted various clinical tasks, such as screening or triage (n=16), disease diagnosis (n=16), risk analysis (n=14), and treatment (n=7). The most commonly addressed diseases and conditions were sepsis (n=6), breast cancer (n=5), diabetic retinopathy (n=4), and polyp and adenoma (n=4). Regarding the evaluation outcomes, we found that 26 studies examined the performance of AI applications in clinical settings, 33 studies examined the effect of AI applications on clinician outcomes, 14 studies examined the effect on patient outcomes, and one study examined the economic impact associated with AI implementation.

Conclusions: This review indicates that research on the clinical implementation of AI applications is still at an early stage despite the great potential. More research needs to assess the benefits and challenges associated with clinical AI applications through a more rigorous methodology.

(*J Med Internet Res* 2021;23(4):e25759) doi: [10.2196/25759](https://doi.org/10.2196/25759)

KEYWORDS

artificial intelligence; machine learning; deep learning; system implementation; clinical practice; review

Introduction

Background

Artificial intelligence (AI) has greatly expanded in health care in the past decade. In particular, AI applications have been applied to uncover information from clinical data and assist health care providers in a wide range of clinical tasks, such as disease diagnosis, triage or screening, risk analysis, and surgical operations [1-4]. According to Accenture analysis, the global health AI market is expected to reach US \$6.6 billion by 2021 and has the potential to grow more than 10 times in the next 5 years [5].

The term “AI” was coined by McCarthy in the 1950s and refers to a branch of computer science wherein algorithms are developed to emulate human cognitive functions, such as learning, reasoning, and problem solving [6]. It is a broadly encompassing term that includes, but is not limited to, machine learning (ML), deep learning (DL), natural language processing (NLP), and computer vision (CV).

Researchers have devoted a great deal of effort to the development of health care AI applications. The number of related articles in the Google Scholar database has grown exponentially since 2000. However, their implementation in real-life clinical practice is not widespread [1,7]. Several reasons may account for this research-practice gap. Specifically, AI algorithms may be subject to technical issues, such as data set shift, overfitting, bias, and lack of generalizability [8], limiting the safe translation of AI research into clinical practice. Further, practical implementation of AI applications can be incredibly challenging. Examples of key challenges that need to be addressed include data sharing and privacy issues, lack of algorithm transparency, the changing nature of health care work, financial concerns, and the demanding regulatory environment [1,3,9-13]. However, the huge potential of health care AI applications can only be realized when they have been integrated into clinical routine workflows.

Research Gap

To the best of our knowledge, this review is the first to systematically examine the role of AI applications in real-life clinical environments. We note that many reviews have been carried out in the area of health care AI. One stream of reviews provided an overview of the current status of AI technology in specific clinical domains, such as breast cancer diagnosis [14], melanoma diagnosis [15], pulmonary tuberculosis diagnosis [16], stroke diagnosis and prediction [17], and diabetes management [18]. Another stream of reviews focused on comparing clinician performance and AI performance to provide the evidence base needed for AI implementation [19-21]. In contrast, our work differs from previous reviews in at least three aspects. First, we review clinical AI applications that provide decision support more broadly and hence do not restrict our scope to a specific clinical domain. Second, we focus on studies that reported the evaluation of clinical AI applications in the real world. We hence exclude studies that discussed the development and validation of clinical AI algorithms without actual implementation. Finally, we report a wide range of evaluation outcomes associated with AI implementation, such

as performance comparison, clinician and patient outcomes, and economic impact.

On the other hand, we note that several viewpoint articles have provided a general outlook of health care AI [1-3,7,9,22]. These articles mainly provided insights into the current status of health care AI and selected a few clinical AI applications as illustrative examples. They might have also discussed the challenges associated with the practical implementation of AI. However, these articles did not discuss the progress of AI implementation that had been made in detail. In contrast, our work aims to provide a comprehensive map of the literature on the evaluation of AI applications in real-life clinical settings. By doing so, we summarize empirical evidence of the benefits and challenges associated with AI implementation and provide suggestions for future research in this important and promising area.

Objective

The objective of this systematic review was to identify and summarize the existing research on AI applications that have been implemented in real-life clinical practice. This helps us better understand the benefits and challenges associated with AI implementation in routine care settings, such as augmenting clinical decision-making capacity, improving care processes and patient outcomes, and reducing health care costs. Specifically, we synthesize relevant studies based on (1) study characteristics, (2) AI application characteristics, and (3) evaluation outcomes and key findings. Considering the research-practice gap, we also provide suggestions for future research that examines and assesses the implementation of AI in clinical practice.

Methods

Search Strategies

The systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [23]. We searched PubMed, Embase, Cochrane Central, and CINAHL in June 2020 to identify relevant articles on AI applications that had been implemented in clinical practice. We limited our search to English-written peer-reviewed journal articles published between January 2010 and May 2020. We chose 2010 as the start period because health care AI research has since taken off.

We used two groups of keywords to identify terms in the titles, abstracts, and keywords of the publications. The first group of keywords had AI-related terms, including “artificial intelligence,” “machine learning,” and “deep learning.” It is worth noting here that AI is a broadly encompassing term and also includes specific AI techniques, such as neural networks, support vector machines, decision trees, and NLP. However, studies using these techniques are highly likely to use “artificial intelligence” or “machine learning” in abstracts or keywords [24]. The second group of keywords had terms related to clinical implementation, including “clinical,” “health,” “healthcare,” “medical,” “implement,” “implementation,” “deploy,” “deployment,” and “adoption.” Details of the search strategy can be found in [Multimedia Appendix 1](#).

Eligibility Criteria

We downloaded and imported all of the identified articles using EndNote X9 (Thomson Reuters) for citation management. After removing duplicates, two researchers (JY and KYN) independently screened the titles and abstracts of the identified articles to determine their eligibility. Disagreements were resolved by discussion between the authors until consensus was reached. The inclusion criteria were as follows: (1) the study implemented an AI application with patients or health care providers in a real-life clinical setting and (2) the AI application provided decision support by emulating clinical decision-making processes of health care providers (eg, medical image interpretation and clinical risk assessment). Medical hardware devices, such as X-ray machines, ultrasound machines, surgery robots, and rehabilitation robots, were outside our scope.

The exclusion criteria were as follows: (1) the study discussed the development and validation of clinical AI algorithms without actual implementation; (2) the AI application provided automation (eg, automated insulin delivery and monitoring) rather than decision support; and (3) the AI application targeted nonclinical tasks, such as biomedical research, operational tasks, and epidemiological tasks. We also excluded conference abstracts, reviews, commentaries, simulation papers, and ongoing studies.

Data Extraction and Charting

Following article selection, we created a data-charting form to extract information from the included articles in the following aspects: (1) study characteristics, (2) AI application characteristics, and (3) evaluation outcomes and key findings ([Textbox 1](#)).

Textbox 1. Components of the data-charting form.

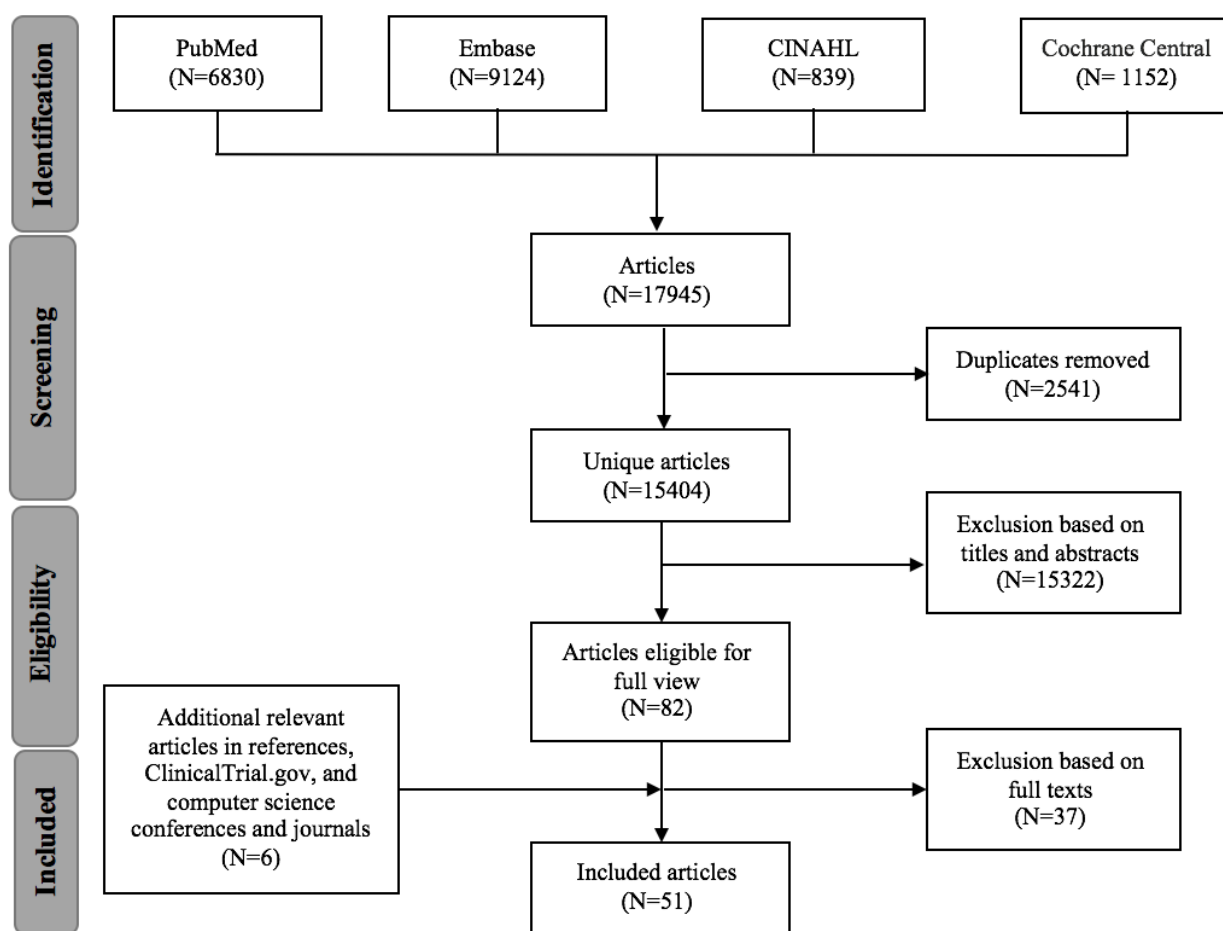
<p>Study characteristics</p> <ul style="list-style-type: none"> • Author, year • Study design • Involved patient(s) and health care provider(s) • Involved hospital(s) and country of the study <p>Artificial intelligence (AI) application characteristics</p> <ul style="list-style-type: none"> • Application description • AI techniques used (eg, neural networks, random forests, and natural language processing) • Targeted clinical tasks • Targeted disease domains and conditions <p>Evaluation outcomes and key findings</p> <ul style="list-style-type: none"> • Performance of AI applications • Clinician outcomes • Patient outcomes • Cost-effectiveness
--

Results

Overview

Our initial search in June 2020 returned a total of 17,945 journal articles (6830 from PubMed, 9124 from Embase, 839 from CINAHL, and 1152 from Cochrane Central) ([Figure 1](#)). We first identified and excluded 2541 duplicates. After that, we

excluded 15,322 articles after screening the titles and abstracts. Thus, 82 articles remained for full-text review, of which 45 were included in this review. Additionally, we identified six relevant articles by examining the references of the included articles, browsing through ClinicalTrial.gov using AI-related keywords, and hand searching premier computer science journals and conferences in AI ([Multimedia Appendix 1](#)). Finally, a total of 51 articles met our inclusion criteria.

Figure 1. Flow diagram of the literature search based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Study Characteristics

Table 1 summarizes the authors, year of publication, study design, involved patients and health care providers, and involved hospitals [25-75]. As shown in **Figure 2**, there was a rising trend in the number of included studies in the last decade, with a recent peak in 2019, suggesting accelerated research activity in this area.

Regarding study design, the 51 studies included 20 observational studies (17 prospective studies and three retrospective studies), 13 randomized controlled trials (RCTs), eight experimental studies, four before-and-after studies, three surveys, one randomized crossover trial, one nonrandomized trial, and one structured interview. It is important to note that observational studies can be categorized into prospective and retrospective studies based on the timing of data collection. In prospective studies, researchers design the research and plan the data collection procedures before any of the subjects have the disease

or develop other outcomes of interest. In retrospective studies, researchers collect existing data on current and past subjects, that is, subjects may have the disease or develop other outcomes of interest before researchers initiate research design and data collection.

Of the 51 studies, 29 (57%) explicitly mentioned the involved patients, two of which had a sample size smaller than 30. On the other hand, 28 (55%) studies provided information about the involved health care providers, of which 17 studies had 10 or fewer providers.

Additionally, 46 (90%) studies mentioned the involved hospitals or clinics (**Figure 3**). Of these, 36 studies were conducted in developed countries, with 20 conducted in the United States, five in the United Kingdom, two each in Australia, Canada, and Japan, one each in Germany, Israel, Spain, and the Netherlands, and one in the United States and South Korea. On the contrary, 10 studies were conducted in developing countries, with eight conducted in China, one in India, and one in India and Kenya.

Table 1. Characteristics of the included studies.

Author, year	Study design	Sample characteristics	Hospital (country)	Evaluation outcomes
Abràmoff et al, 2018 [25]	Observational study (prospective)	819 patients	10 primary care clinics (United States)	AP ^a (sensitivity, specificity, imageability rate)
Aoki et al, 2020 [26]	Experimental study (crossover design)	6 physicians	The University of Tokyo Hospital (Japan)	CO ^b (reading time, mucosal break detection rate)
Arbabshirani et al, 2018 [27]	Observational study (prospective)	347 routine head CT ^c scans of patients	Geisinger Health System (United States)	AP (AUC ^d , accuracy, sensitivity, specificity) CO (time to diagnosis)
Bailey et al, 2013 [28]	Crossover RCT ^e	20,031 patients	Barnes-Jewish Hospital (United States)	PO ^f (ICU ^g transfer, hospital mortality, hospital LOS ^h)
Barinov et al, 2019 [29]	Experiment (within subjects)	3 radiologists	NR ⁱ	AP (AUC) CO (diagnostic accuracy)
Beaudoin et al, 2016 [30]	Observational study (prospective)	350 patients (515 prescriptions)	Centre hospitalier universitaire de Sherbrooke (Canada)	AP (number of triggered recommendations, precision, recall, accuracy)
Bien et al, 2018 [31]	Experimental study (within subjects)	9 clinical experts	Stanford University Medical Center (United States)	AP (AUC) CO (specificity, sensitivity, accuracy)
Brennan et al, 2019 [32]	Nonrandomized trial	20 physicians	An academic quaternary care institution (United States)	AP (AUC) CO (risk assessment changes, AUC, usability)
Chen et al, 2020 [33]	RCT	437 patients	Renmin Hospital, Wuhan University (China)	CO (blind spot rate)
Connell et al, 2019 [34]	Before-after study	2642 patients	Royal Free Hospital, Barnet General Hospital (United Kingdom)	PO (renal recovery rate, other clinical outcomes, care process)
Eshel et al, 2017 [35]	Observational study (prospective)	6 expert microscopists	Apollo Hospital, Chennai (India); Aga Khan University Hospital (Kenya)	AP (sensitivity, specificity, species identification accuracy, device parasite count)
Giannini et al, 2019 [36]	Before-after study	22,280 patients in the silent period, 32,184 patients in the alert period	3 urban acute hospitals under University of Pennsylvania Health System (United States)	AP (sensitivity, specificity) PO (mortality, discharge disposition, ICU transfer, time to ICU transfer, clinical process measures)
Ginestra et al, 2019 [37]	Survey	43 nurses and 44 health care providers	A tertiary teaching hospital in Philadelphia (United States)	CO (nurse and provider perceptions)
Gómez-Vallejo et al, 2016 [38]	Observational study (retrospective)	1800 patients (2569 samples)	A Spanish National Health System hospital (Spain)	AP (accuracy) CO (system perceptions)
Grunwald et al, 2016 [39]	Observational study (retrospective)	15 patients, 3 neuroradiologists	A comprehensive stroke center (Germany)	AP (e-ASPECTS performance)
Kanagasingam et al, 2018 [40]	Observational study (prospective)	193 patients, 4 physicians	A primary care practice in Midland (Australia)	AP (sensitivity, specificity, PPV ^j , NPV ^k)
Keel et al, 2018 [41]	Survey	96 patients	St Vincent's Hospital, University Hospital Geelong (Australia)	AP (sensitivity and specificity, assessment time) PO (patient acceptability)
Kiani et al, 2020 [42]	Experimental study (within subjects)	11 pathologists	Stanford University Medical Center (United Kingdom)	AP (accuracy) CO (accuracy)
Lagani et al, 2015 [43]	Observational study (prospective)	2 health care providers	Chorleywood Health Centre (United Kingdom)	AP (system performance) CO (usability)

Author, year	Study design	Sample characteristics	Hospital (country)	Evaluation outcomes
Lin et al, 2019 [44]	RCT	350 patients	5 ophthalmic clinics (China)	AP (accuracy, PPV, NPV) CO (time to diagnosis) PO (patient satisfaction)
Lindsey et al, 2018 [45]	Experimental study (within subjects)	40 practicing emergency clinicians	Hospital for Special Surgery (United States)	AP (AUC) CO (sensitivity, specificity, misinterpretation rate)
Liu et al, 2020 [46]	RCT	1026 patients	No. 988 Hospital of Joint Logistic Support Force of PLA (China)	CO (ADR ¹ , PDR ^m , number of detected adenomas and polyps)
Mango et al, 2020 [47]	Experimental study (within subjects)	15 physicians	13 different medical centers (United States)	AP (AUC, sensitivity, specificity) CO (AUC, interreliability, intrareliability)
Martin et al, 2012 [48]	Observational study (prospective)	214 patients	13 different medical centers (United States)	AP (sensitivity, PPV) PO (ACSC ⁿ , care-supported activities)
McCoy and Das, 2017 [49]	Before-after study	1328 patients	Cape Regional Medical Center (United States)	PO (hospital mortality, hospital LOS, readmission rate)
McNamara et al, 2019 [50]	Observational study (prospective)	3 breast cancer experts	John Theurer Cancer Center (United States)	CO (decision making)
Mori et al, 2018 [51]	Observational study (prospective)	791 patients, 23 endoscopists	Showa University Northern Yokohama Hospital (Japan)	AP (NPV) CO (time to diagnosis)
Nagaratnam et al, 2020 [52]	Observational study (retrospective)	1 patient	Royal Berkshire Hospital (United Kingdom)	PO (patient care and clinical outcomes)
Natarajan et al, 2019 [53]	Observational study (prospective)	213 patients	Dispensaries under Municipal Corporation of Greater Mumbai (India)	AP (sensitivity, specificity)
Nicolae et al, 2020 [54]	RCT	41 patients	Sunnybrook Odette Cancer Centre (Canada)	AP (day 30 dosimetry) CO (planning time)
Park et al, 2019 [55]	Experimental study (within subjects)	8 clinicians	Stanford University Medical Center (United States)	CO (specificity, sensitivity, accuracy interrater agreement, time to diagnosis)
Romero-Brufau et al, 2020 [56]	Pre-post survey	81 clinical staff	3 primary-care clinics in Southwest Wisconsin (United States)	CO (attitudes about AI ^o in the workplace)
Rostill et al, 2018 [57]	RCT	204 patients, 204 caregivers	NHS, Surrey and Hampshire (United Kingdom)	CO (system evaluations) PO (early clinical interventions, patient evaluations)
Segal et al, 2014 [58]	Observational study (prospective)	16 pediatric neurologists	Boston Children's Hospital (United States)	CO (diagnostic errors, diagnosis relevance, number of workup items)
Segal et al, 2016 [59]	Observational study (prospective)	26 clinicians	Boston Children's Hospital (United States)	CO (diagnostic errors)
Segal et al, 2017 [60]	Structured interviews	10 medical specialists	Geisinger Health System and Intermountain Healthcare (United States)	CO (system perceptions)
Segal et al, 2019 [61]	Observational study (prospective)	3160 patients (315 prescription alerts)	Sheba Medical Center (Israel)	AP (accuracy, clinical validity, and usefulness) PO (changes in medical orders)
Shimabukuro et al, 2017 [62]	RCT	142 patients	University of California San Francisco Medical Center (United States)	PO (LOS, in-hospital mortality)

Author, year	Study design	Sample characteristics	Hospital (country)	Evaluation outcomes
Sim et al, 2020 [63]	Observational study (prospective)	12 radiologists	4 medical centers (United States and South Korea)	AP (sensitivity, FPPI) CO (sensitivity, FPPI, decision change)
Steiner et al, 2018 [64]	Experimental study (within subjects)	6 anatomic pathologists	NR	CO (sensitivity, average review per image, interpretation difficulty)
Su et al, 2020 [65]	RCT	623 patients, 6 endoscopists	Qilu Hospital of Shandong University (China)	CO (ADR, PDR, number of adenomas and polyps, withdrawal time, adequate bowel preparation rate)
Titano et al, 2018 [66]	RCT	2 radiologists	NR	CO (time to diagnosis, queue of urgent cases)
Vandenberghe et al, 2017 [67]	Observational study (prospective)	1 pathologist and 2 HER2 raters	NR	CO (decision concordance, decision modification)
Voerman et al, 2019 [68]	Before-after study	NR	Five Rivers Medical Center, Pocahontas (United States)	CE ^q (average total costs per patient) PO (numbers of patients with clostridium difficile and antibiotic-resistant infections, LOS, antibiotic use)
Wang et al, 2019 [69]	RCT	1058 patients, 8 physicians	Sichuan Provincial People's Hospital (China)	CO (ADR, PDR, number of adenomas per patient)
Wang et al, 2019 [70]	RCT	75 patients	4 primary care clinics affiliated with Brigham and Women's Hospital (United States)	CO (anticoagulation prescriptions)
Wang et al, 2020 [71]	RCT	962 patients	Caotang branch hospital of Sichuan Provincial People's Hospital (China)	CO (ADR, PDR, number of adenomas and polyps per colonoscopy)
Wijnberge et al, 2020 [72]	RCT	68 patients	Amsterdam UMC (Netherlands)	PO (median time-weighted average of hypotension, median time of hypotension, treatment, time to intervention, adverse events)
Wu et al, 2019 [73]	Observational study (prospective)	3600 residents	3 ophthalmologists, community healthcare centers (China)	AP (AUC) CO (ophthalmologist-to-population service ratio)
Wu et al, 2019 [74]	RCT	303 patients, 6 endoscopists	Renmin hospital of Wuhan University (China)	AP (accuracy, completeness of photo documentation) CO (blind spot rate, number of ignored patients, inspection time) PO (adverse events)

Author, year	Study design	Sample characteristics	Hospital (country)	Evaluation outcomes
Yoo et al, 2018 [75]	Observational study (prospective)	50 patients, 1 radiologist	NR (Korea)	AP (sensitivity, specificity, PPV, NPV, accuracy) CO (sensitivity, specificity, PPV, NPV, accuracy)

^aAP: application performance.

^bCO: clinician outcomes.

^cCT: computed tomography.

^dAUC: area under the curve.

^eRCT: randomized controlled trial.

^fPO: patient outcomes.

^gICU: intensive care unit.

^hLOS: length of stay.

ⁱNR: not reported.

^jPPV: positive-predictive value.

^kNPV: negative-predictive value.

^lADR: adenoma detection rate.

^mPDR: polyp detection rate.

ⁿACSC: ambulatory care sensitive admissions.

^oAI: artificial intelligence.

^pFFPI: false-positive per image.

^qCE: cost-effectiveness.

Figure 2. Distribution of the included articles from 2010 to 2020.

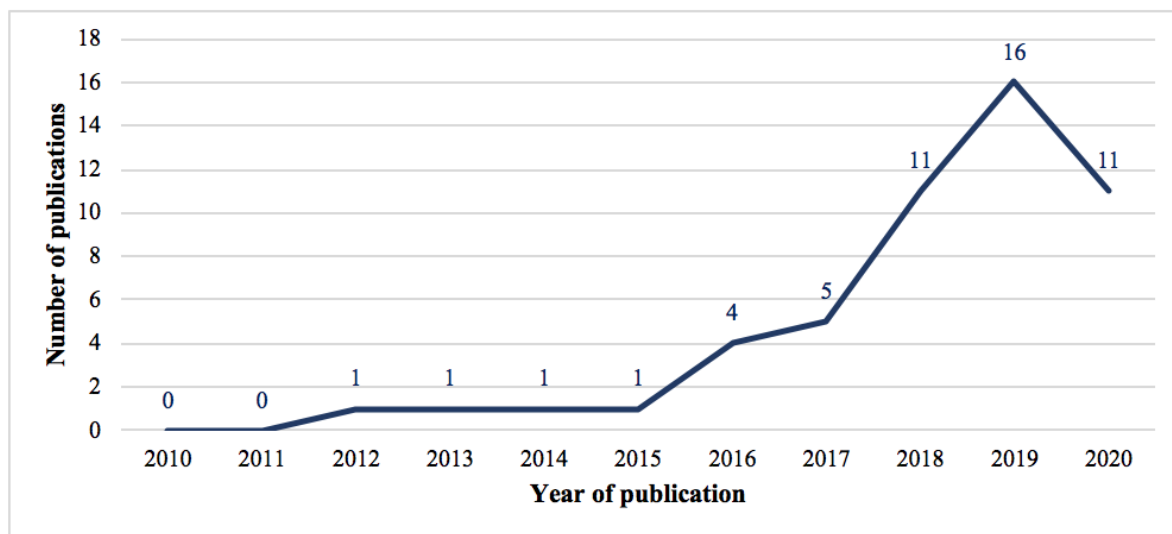
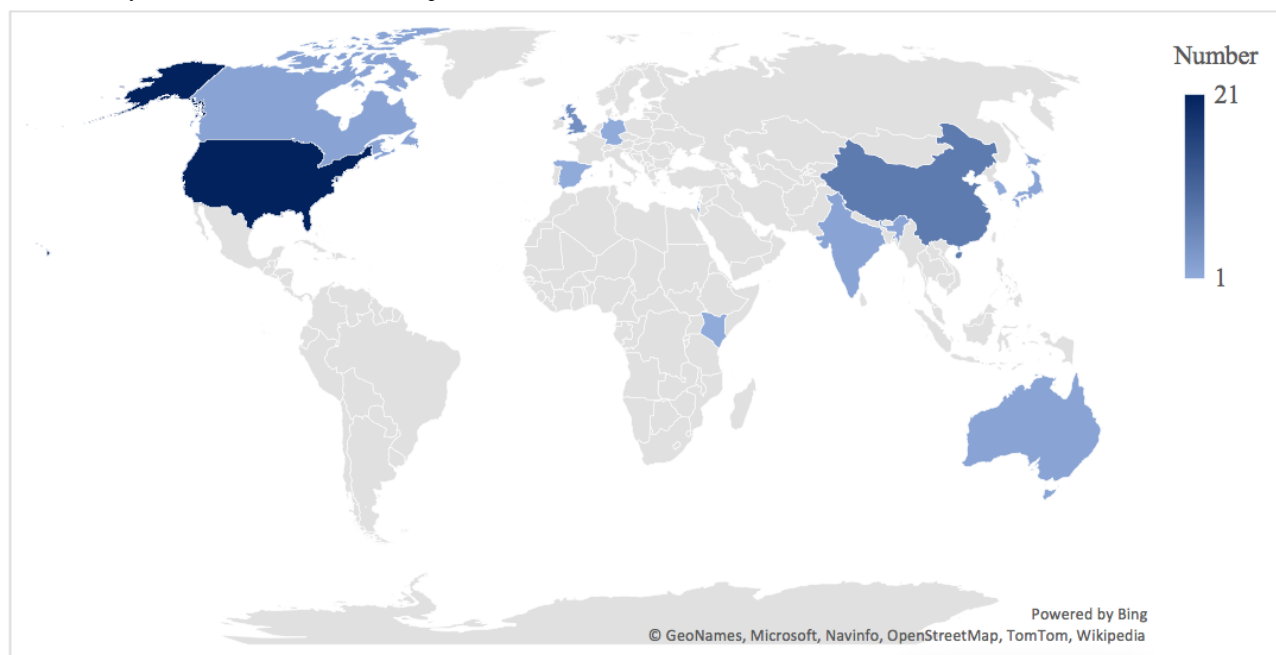


Figure 3. Country distribution of the involved hospitals.

Quality Assessment

Considering the heterogeneity of study types included in the review, we only assessed the risk of bias of 13 RCTs using the Cochrane Collaboration Risk of Bias tool ([Multimedia Appendix 2](#)) [76]. Overall, the total score of the RCTs ranged from 0 [57] to 6 [44,65,69], with a mean value of 3.84. Specifically, eight studies reported random sequence generation [33,44,62,65,69,71,72], and three studies explicitly stated that the allocation was concealed [62,65,72]. Only two studies were double blinded [66,71]. Blinding of participants was unsuccessful in two studies [62,72] and was unclear in six studies [44,46,54,57,69,70]. Blinding of outcome assessment was unsuccessful in seven studies [33,46,62,65,69,70,74] and was unclear in one study [57]. Three studies did not clearly state whether they had complete data for the enrolled participants [57,66,77]. All of the 13 studies had a low risk of selective reporting bias. Other potential sources of bias included a small sample size [62,70,72], a short study period [62], and a lack of detailed information regarding RCTs and follow-ups [57,66].

AI Application Characteristics

Among the 51 studies, two did not disclose any information regarding the AI techniques used. Among the remaining 49 studies, the most popular ML technique was neural networks ($n=22$), followed by random forests ($n=3$), Bayesian pattern matching ($n=3$), support vector machine ($n=2$), decision tree ($n=2$), and deep reinforcement learning ($n=2$). We also found that the included AI applications mainly provided decision support in the following four categories of clinical tasks: disease screening or triage ($n=16$), disease diagnosis ($n=16$), risk analysis ($n=14$), and treatment ($n=7$). Further, AI applications in 46 (94%) studies targeted one or more specific diseases and conditions. The most prevalent diseases and conditions were sepsis ($n=6$), breast cancer ($n=5$), diabetic retinopathy ($n=4$), polyp and adenoma ($n=4$), cataracts ($n=2$), and stroke ($n=2$).

Details of AI application characteristics are provided in [Multimedia Appendix 3](#).

Evaluation Outcomes

We categorized the evaluation outcomes in our review studies into the following four types: performance of AI applications, clinician outcomes, patient outcomes, and cost-effectiveness, as can be seen in [Table 1](#) and [Multimedia Appendix 4](#).

Performance of AI Applications

Twenty-six studies evaluated the performance of AI applications in real-life clinical settings [25,27,29-32,35,36,38-43,45,47,48,51,52,54,61,63,73-75,78]. Commonly used performance metrics included accuracy, area under the curve (AUC), specificity, sensitivity, positive-predictive value (PPV), and negative-predictive value (NPV). Of these, 24 studies reported acceptable and satisfactory performance of AI applications in practice. For example, one study [25] conducted a pivotal trial of the IDx-DR diagnostic system (IDx, LLC) to detect diabetic retinopathy in 10 primary clinic offices in the United States. They reported that IDx-DR had a sensitivity of 87.2%, a specificity of 90.7%, and an imageability rate of 96.1%, exceeding prespecified endpoints. Based on the results, IDx-DR became the first Food and Drug Administration (FDA)-authorized AI diagnostic system, with the potential to improve early detection of diabetic retinopathy and prevent vision loss in thousands of patients with diabetes.

On the contrary, two studies found that AI applications failed to outperform health care providers and needed further improvement [40,44]. In particular, one RCT [44] examined the performance of CC-Cruiser, an AI-based platform for childhood cataracts, in five ophthalmic clinics in China. The authors found that CC-Cruiser had considerably lower accuracy, PPV, and NPV than senior consultants in diagnosing childhood congenital cataracts and making treatment decisions. Another study [40] evaluated the performance of an AI-based diabetic retinopathy grading system in a primary care office in Australia

and found that the AI system had a high false-positive rate with a PPV of 12%. Specifically, of the 193 patients who consented to the study, the AI system identified 17 patients with severe diabetic retinopathy that required referral. However, only two patients were correctly identified, and the remaining 15 patients were false positives.

Clinician Outcomes

Thirty-three studies examined the effect of AI applications on clinician outcomes, that is, clinician decision making, clinician workflow and efficiency, and clinician evaluations and acceptance of AI applications [26,27,29,31-33,37,38,42-47,50,51,54-60,64,65,67,69-71,73-75].

AI applications have the potential to provide clinical decision support. From our review, 16 studies demonstrated that AI applications could enhance clinical decision-making capacity [31-33,45-47,50,55,58,59,63,64,67,69,71,74,75]. For example, Brennan et al [32] found that clinicians gained knowledge after interacting with MySurgery, an algorithm for preoperative risk assessments, and improved their risk assessment performance as a result. On the contrary, two studies did not find any evidence for enhanced decision-making [26,42]. One possible explanation is that AI may provide misleading recommendations, offsetting the benefits of AI. Specifically, Kiani et al [42] evaluated the effect of a DL-based system for live cancer classification on the diagnostic performance of 11 pathologists and found that AI use did not greatly improve the diagnostic accuracy. They further noted that AI improved accuracy when it provided correct predictions and harmed accuracy when it provided wrong predictions. Aoki et al [26] examined the impact of a DL-based system for mucosal break detection on endoscopists in reading small bowel capsule endoscopy. They found that the system failed to improve the mucosal break detection performance of endoscopists, particularly trainees.

Seven studies were aimed at clinician workflow and efficiency [26,27,44,51,54,66,73]. Of these, six studies found that AI accelerated the time needed for clinical tasks and improved the existing workflow [26,27,44,51,54,66]. For example, Titano et al [66] found that a DL-based cranial image triage algorithm processed and interpreted images 150 times faster than human radiologists (1.2 seconds vs 177 seconds) and appropriately escalated urgent cases, enhancing the triage of cases in the radiology workflow. The only exception is the work of Wu et al [74], which assessed the quality improvement system WISENSE for blind-spot monitoring and procedure timing during esophagogastroduodenoscopy. This study found that WISENSE helped endoscopists monitor and control their time on each procedure and increased inspection time as a result.

Finally, clinician perceptions and acceptance of AI applications were examined in seven studies [32,37,38,43,56,57,60]. Particularly, five out of the seven studies reported overall positive perceptions of AI applications [32,38,43,57,60]. For example, Brennan et al [32] asked 20 surgical intensivists to use and evaluate MySurgeryRisk for preoperative risk prediction in a simulated clinical workflow. Most respondents indicated that MySurgeryRisk was useful and easy to use and believed that it would be helpful for decision making. On the other hand,

the remaining two studies reported mixed or even negative evaluations of AI [37,56]. Specifically, Ginestra et al [37] assessed physician evaluations of an ML-based sepsis prediction system in a tertiary teaching hospital and found that only 16% of health care providers perceived system-generated sepsis alerts to be helpful. The negative evaluations could be attributed to providers' low confidence in alerts, low algorithm transparency, and a lack of established actions after alerts. Romero-Brufau et al [56] reported survey results from implementing an AI-based clinical decision support system in a regional health system practice and found that only 14% of clinical staff were willing to recommend the system. Staff feedback revealed that some system-recommended interventions were inadequate and inappropriate.

Patient Outcomes

Fourteen studies reported patient outcomes [28,34,36,41,44,48,49,52,57,61,62,68,72,74]. In 11 of the 14 studies, researchers examined the effect of AI on clinical processes and outcomes, such as hospital length of stay, in-hospital mortality, intensive care unit (ICU) transfer, readmission, and time to intervention [28,34,36,48,49,52,57,61,62,68,72,74]. The results were inconsistent. Most studies reported improved clinical outcomes (n=8) [36,48,52,57,61,62,68,72,74]. For example, one RCT [62] implemented and assessed an ML-based severe sepsis prediction algorithm (Dascena) in two ICUs at the University of California San Francisco Medical Center. They found that the algorithm implementation greatly decreased the hospital length of stay from 13.0 days to 10.3 days and decreased the in-hospital mortality rate from 21.3% to 8.96%. However, three of the studies did not find evidence for improved clinical outcomes, indicating the limited applicability of the algorithms in their current form [28,34,36]. In particular, Bailey et al [28] examined the effect of an ML-based algorithm that generated real-time alerts for clinical deterioration in hospitalized patients. They found that providing alerts alone could not reduce the hospital length of stay and the in-hospital mortality. Connell et al [34] examined the effect of a novel digitally enabled care pathway for acute kidney injury management and found no step changes in the renal recovery rate and other secondary clinical outcomes following the intervention. Giannini et al [36] developed and implemented a sepsis prediction algorithm in a tertiary teaching hospital system. The results showed that the algorithm-generated alerts had a limited impact on clinical processes and could not reduce mortality, discharge dispositions, or transfer to the ICU. Future algorithm optimization is thus needed.

Three studies examined how patients evaluated AI applications, and all of them reported positive results [41,44,57]. Keel et al [41] evaluated patient acceptability of an AI-based diabetic retinopathy screening tool in an endocrinology outpatient setting. They found that 96% (92/96) of the screened patients were satisfied with the AI tool and 78% (43/55) of the patients in the follow-up survey preferred AI screening over manual screening, suggesting that the AI tool was well-accepted by patients. Lin et al [44] assessed patient satisfaction with CC-Cruiser for childhood cataracts and found that patients were slightly more satisfied with CC-Cruiser in comparison with senior consultants. One explanation is that childhood cataracts may cause

irreversible vision impairment and even blindness without early intervention. Therefore, parents of patients appreciated the faster diagnosis of CC-Cruiser. Rostill et al [57] assessed an Internet of Things (IoT) system for dementia care and found that dementia patients trusted the system and would like to recommend it.

Cost-Effectiveness

The economic impact of AI implementation in clinical practice was addressed in only one study [68]. This study reported that the implementation of an ML-based system for antibiotic stewardship reduced costs by US \$25,611 for sepsis and US \$3630 for lower respiratory tract infections compared with usual care.

Discussion

Principal Findings

AI applications have huge potential to augment clinician decision making, improve clinical care processes and patient outcomes, and reduce health care costs. Our review seeks to identify and summarize the existing studies on AI applications that have been implemented in real-life clinical practice. It yields the following interesting findings.

First, we note that the number of included studies was surprisingly small considering the tremendous number of studies on health care AI. In particular, most of the health care AI studies were proof-of-concept studies that focused on AI algorithm development and validation using retrospective clinical data sets. In contrast, only a handful of studies implemented and evaluated AI in a clinical environment. To ensure safe adoption, however, an AI application should provide solid scientific evidence for its effectiveness relative to the standard of care. Therefore, we urge the health care AI research community to work closely with health care providers and institutions to demonstrate the potential of AI in real-life clinical settings.

Second, more than two-thirds of the included articles were from developed economies, of which more than half were from the United States, suggesting that developed countries are at the forefront of health care AI development and deployment. This is consistent with the fact that top health AI companies and start-ups (eg, Google Health, IBM Watson Health, and Babylon Health) are mainly located in the United States and Europe. This finding should be interpreted with caution because we excluded non-English-written articles, even though our search had identified 890 non-English publications. We did not include these non-English articles because it is difficult to conduct an unbiased analysis owing to translation difficulty and variation. The imbalanced distribution of articles by country or economic development status could be attributed to the fact that researchers from low-income countries have a very low publication rate.

However, it is worth noting that 8 (16%) of our articles were from China, suggesting that China has been extensively applying health care AI and conducting health care AI research. Indeed, hospitals, technology companies, and the Chinese government have been driving clinical AI deployment with the aim to

alleviate doctor shortages, relieve medical resource inequality, and reduce health care costs [79-82], and Chinese researchers have acquired the capability to publish in international English journals.

Third, the quality of research on clinical AI evaluation needs to be improved in the future. Our review revealed that only 13 (26%) studies were RCTs and most of them suffered from moderate to high risk of bias. Eight studies were experimental studies, and all of them adopted a cross-over design or within-subjects design and were hence susceptible to confounding effects. With respect to sample information, only 8 (16%) studies provided information on both patients and health care providers, and 14 (28%) studies used a sample size smaller than 20 (Table 1), limiting the generalizability of their results. Regarding the evaluation design, one-third of the studies (n=17, 33%) did not include a comparison group (Multimedia Appendix 4), limiting the ability to identify the added value of AI applications compared with the current best practice. Given that health care providers may hold different perceptions toward different AI systems of varying performance and reliability, it would be helpful if the studies provide a transparent description of the AI system's architecture, accuracy or reliability performance, and possible risks. Unfortunately, in our review, 21 studies did not provide adequate information about the architecture of the AI applications [25,29,32,34,37,44,46,50,52,54,56-62,65,68,70,75] and 22 studies did not reveal the performance and possible risks of AI under evaluation [26,29,34,37,39,46,48-50,52,54,56-62,64,65,68,69]. Further, considering that some self-evolving adaptive clinical AI applications continuously incorporate the latest clinical practice data and published evidence, it is important to undertake periodic monitoring and recalibration of AI applications to ensure that they are working as expected. Finally, we found that more than half of the studies (n=29, 57%) investigated only one aspect of evaluation outcome (Multimedia Appendix 4). We encourage future research to conduct a more comprehensive assessment of the quality of clinical AI applications as well as their impacts on clinicians, patients, and health care institutions. This will facilitate the comparison and selection of alternative AI solutions in the same clinical domain.

Fourth, our analysis indicated that AI applications could provide effective decision support, albeit in certain contexts. For instance, the augmenting role of AI in clinical decision-making capacity can be affected by the level of expertise. In particular, two studies suggested that junior physicians were more likely to benefit from AI than senior physicians because they had a higher tendency to reconsider and modify their clinical decisions when encountering disconfirming AI suggestions [38,47]. However, it is worth noting that AI can be misleading sometimes. For example, one study from our review speculated that trainee endoscopists may feel confused about false-positive results from an AI screening tool owing to limited reading experience and, as a result, ignore AI-marked lesions of small-bowel mucosal breaks [26]. It is therefore important for future research to examine under what circumstances physicians could benefit more from AI applications. However, we are sanguine that when AI technology is sufficiently mature and

accurate to become the evidence-based best practice, its use would become part of routine clinical care in the future.

With respect to AI acceptance, we observed that health care providers expressed negative feelings toward AI in two studies [37,56], indicating that barriers existed in the incorporation of AI into the routine workflow. However, an elaboration of AI implementation barriers will be lengthy and is beyond the scope of this work, and we refer interested readers to the reports by Kelly et al [8], Ngiam and Khor [83], Lysaght et al [84], Shaw et al [10], and Yu and Kohane [12] for more details.

Fifth, for most of the included studies on patient outcomes, we found that they did not examine the clinical processes and interventions in detail. However, AI applications without appropriate and useful interventions may be ineffective at improving patient outcomes. For example, Bailey et al [28] found that simply notifying the nursing staff of clinical deterioration risks was not able to improve the outcomes of high-risk patients. Effective patient-specific interventions are needed. Therefore, future research may design and evaluate patient-directed interventions to enhance the clinical effectiveness of AI applications.

Moreover, three of the included studies suggested that patients and their families were highly satisfied with health care AI owing to its convenience and efficiency [41,44,57]. However, this may not always be the case. Prior research has shown that patients preferred to receive primary care from a human provider than AI even if the care from the health provider entailed a higher misdiagnosis risk [85]. The reason is that they perceived AI to be less capable in considering their unique circumstances. Additionally, patients may disparage physicians aided by a clinical decision support system and perceive them as less capable and professional than their unaided counterparts [86]. Further studies to explore the possible patient concerns and resistance toward health AI applications should be considered.

Finally, according to an Accenture survey, more than half of health care institutions are optimistic that AI will reduce costs and improve revenue despite the high initial costs associated with AI implementation [87]. However, only one included study documented the economic outcomes of AI implementation. This highlights the need to conduct more cost-effectiveness analyses of AI applications in clinical practice.

Acknowledgments

This study was supported by NSCP (grant no. N-171-000-499-001) from the National University of Singapore. It was also supported by Dean Strategic Fund-Health Informatics (grant no. C-251-000-068-001) from the National University of Singapore.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[\[PDF File \(Adobe PDF File\), 86 KB-Multimedia Appendix 1\]](#)

Limitations

This review has several limitations. First, we only included peer-reviewed English-written journal articles. It is plausible that some relevant articles were written in other languages or published in conferences, workshops, and news reports. As noted earlier, this may partly explain the imbalanced country distribution of the reviewed articles. Moreover, we did not include articles that were published before 2010 because AI only started to make in-roads in the clinical field in the last decade, as evident in our search results. Moreover, we only reviewed premium computer science conferences and journals without comprehensively examining engineering and computer science databases. This should be less of a concern here because we found that computer science conferences and journals mainly focus on the training and validation of novel AI algorithms without actual deployment. Still, future research can expand the search scope to gain deeper insights into state-of-the-art clinical AI algorithms.

Another concern is that some AI applications may have been implemented in real-world clinical practice without any openly accessible publications. For example, IDx-DR, the first FDA-approved AI system, has been implemented in more than 20 health care institutions such as University of Iowa Health Care [88]. However, our search only identified one related published result [25]. Clinical practitioners should take a more active role in reporting AI evaluation and use results in their daily practice in the future.

Conclusions

AI applications have tremendous potential to improve patient outcomes and improve care processes. Based on the literature presented in this review, there is great interest to develop AI tools to support clinical workflows, with increasing high-quality evidence being generated. However, there is currently insufficient level 1 evidence to advocate the routine use of health care AI for decision support, hindering the growth of health care AI and presenting potential risks to patient safety. We thus conclude that it is important to conduct robust RCTs to benchmark AI-aided care processes and outcomes to the current best practice. A rigorous, robust, and comprehensive evaluation of health care AI will help move from theory to clinical practice.

Multimedia Appendix 2

Quality assessments of randomized controlled trials based on the Cochrane Collaboration Risk of Bias Tool.

[[XLSX File \(Microsoft Excel File\), 10 KB-Multimedia Appendix 2](#)]

Multimedia Appendix 3

Artificial intelligence application characteristics.

[[PDF File \(Adobe PDF File\), 112 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

Evaluation outcomes and main results of the included studies.

[[PDF File \(Adobe PDF File\), 153 KB-Multimedia Appendix 4](#)]

References

1. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019 Jan;25(1):30-36 [[FREE Full text](#)] [doi: [10.1038/s41591-018-0307-0](https://doi.org/10.1038/s41591-018-0307-0)] [Medline: [30617336](#)]
2. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017 Dec;2(4):230-243 [[FREE Full text](#)] [doi: [10.1136/svn-2017-000101](https://doi.org/10.1136/svn-2017-000101)] [Medline: [29507784](#)]
3. Yu K, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018 Oct;2(10):719-731 [[FREE Full text](#)] [doi: [10.1038/s41551-018-0305-z](https://doi.org/10.1038/s41551-018-0305-z)] [Medline: [31015651](#)]
4. Triantafyllidis AK, Tsanas A. Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature. *J Med Internet Res* 2019 Apr 05;21(4):e12286 [[FREE Full text](#)] [doi: [10.2196/12286](https://doi.org/10.2196/12286)] [Medline: [30950797](#)]
5. Artificial intelligence (AI): healthcare's new nervous system. Accenture. URL: <https://www.accenture.com/us-en/insight-artificial-intelligence-healthcare%C2%A0> [accessed 2021-04-05]
6. McCarthy J, Minsky M, Rochester N, Shannon C. A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine* 2006;27(4):12. [doi: [10.1609/aimag.v27i4.1904](https://doi.org/10.1609/aimag.v27i4.1904)]
7. Panch T, Mattie H, Celi LA. The "inconvenient truth" about AI in healthcare. *NPJ Digit Med* 2019;2:77 [[FREE Full text](#)] [doi: [10.1038/s41746-019-0155-4](https://doi.org/10.1038/s41746-019-0155-4)] [Medline: [31453372](#)]
8. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019 Oct 29;17(1):195 [[FREE Full text](#)] [doi: [10.1186/s12916-019-1426-2](https://doi.org/10.1186/s12916-019-1426-2)] [Medline: [31665002](#)]
9. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. *NPJ Digit Med* 2019;2:69 [[FREE Full text](#)] [doi: [10.1038/s41746-019-0148-3](https://doi.org/10.1038/s41746-019-0148-3)] [Medline: [31372505](#)]
10. Shaw J, Rudzicz F, Jamieson T, Goldfarb A. Artificial Intelligence and the Implementation Challenge. *J Med Internet Res* 2019 Jul 10;21(7):e13659 [[FREE Full text](#)] [doi: [10.2196/13659](https://doi.org/10.2196/13659)] [Medline: [31293245](#)]
11. Cohen IG, Amarasingham R, Shah A, Xie B, Lo B. The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Aff (Millwood)* 2014 Jul;33(7):1139-1147 [[FREE Full text](#)] [doi: [10.1377/hlthaff.2014.0048](https://doi.org/10.1377/hlthaff.2014.0048)] [Medline: [25006139](#)]
12. Yu K, Kohane IS. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf* 2019 Mar;28(3):238-241. [doi: [10.1136/bmjqs-2018-008551](https://doi.org/10.1136/bmjqs-2018-008551)] [Medline: [30291179](#)]
13. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019 Jan;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](#)]
14. Gardezi SJS, Elazab A, Lei B, Wang T. Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review. *J Med Internet Res* 2019 Jul 26;21(7):e14464 [[FREE Full text](#)] [doi: [10.2196/14464](https://doi.org/10.2196/14464)] [Medline: [31350843](#)]
15. Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. *Br J Dermatol* 2009 Sep;161(3):591-604. [doi: [10.1111/j.1365-2133.2009.09093.x](https://doi.org/10.1111/j.1365-2133.2009.09093.x)] [Medline: [19302072](#)]
16. Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS One* 2019;14(9):e0221339 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0221339](https://doi.org/10.1371/journal.pone.0221339)] [Medline: [31479448](#)]
17. Wang W, Kiik M, Peek N, Curcin V, Marshall IJ, Rudd AG, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS One* 2020;15(6):e0234722 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0234722](https://doi.org/10.1371/journal.pone.0234722)] [Medline: [32530947](#)]
18. Contreras I, Vehi J. Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *J Med Internet Res* 2018 May 30;20(5):e10775 [[FREE Full text](#)] [doi: [10.2196/10775](https://doi.org/10.2196/10775)] [Medline: [29848472](#)]
19. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 2020 Mar 25;368:m689 [[FREE Full text](#)] [doi: [10.1136/bmj.m689](https://doi.org/10.1136/bmj.m689)] [Medline: [32213531](#)]

20. Shen J, Zhang CJP, Jiang B, Chen J, Song J, Liu Z, et al. Artificial Intelligence Versus Clinicians in Disease Diagnosis: Systematic Review. *JMIR Med Inform* 2019 Aug 16;7(3):e10010 [FREE Full text] [doi: [10.2196/10010](https://doi.org/10.2196/10010)] [Medline: [31420959](https://pubmed.ncbi.nlm.nih.gov/31420959/)]
21. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, 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 Oct;1(6):e271-e297 [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)] [Medline: [33323251](https://pubmed.ncbi.nlm.nih.gov/33323251/)]
22. Quer G, Muse ED, Nikzad N, Topol EJ, Steinhubl SR. Augmenting diagnostic vision with AI. *The Lancet* 2017 Jul;390(10091):221 [FREE Full text] [doi: [10.1016/S0140-6736\(17\)31764-6](https://doi.org/10.1016/S0140-6736(17)31764-6)] [Medline: [30078917](https://pubmed.ncbi.nlm.nih.gov/30078917/)]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
24. Wolff J, Pauling J, Keck A, Baumbach J. The Economic Impact of Artificial Intelligence in Health Care: Systematic Review. *J Med Internet Res* 2020 Feb 20;22(2):e16866 [FREE Full text] [doi: [10.2196/16866](https://doi.org/10.2196/16866)] [Medline: [32130134](https://pubmed.ncbi.nlm.nih.gov/32130134/)]
25. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;1:39 [FREE Full text] [doi: [10.1038/s41746-018-0040-6](https://doi.org/10.1038/s41746-018-0040-6)] [Medline: [31304320](https://pubmed.ncbi.nlm.nih.gov/31304320/)]
26. Aoki T, Yamada A, Aoyama K, Saito H, Fujisawa G, Odawara N, et al. Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading. *Dig Endosc* 2020 May;32(4):585-591. [doi: [10.1111/den.13517](https://doi.org/10.1111/den.13517)] [Medline: [31441972](https://pubmed.ncbi.nlm.nih.gov/31441972/)]
27. Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, Suever JD, Geise BD, Patel AA, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *NPJ Digit Med* 2018;1:9 [FREE Full text] [doi: [10.1038/s41746-017-0015-z](https://doi.org/10.1038/s41746-017-0015-z)] [Medline: [31304294](https://pubmed.ncbi.nlm.nih.gov/31304294/)]
28. Bailey TC, Chen Y, Mao Y, Lu C, Hackmann G, Micek ST, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. *J Hosp Med* 2013 May;8(5):236-242. [doi: [10.1002/jhm.2009](https://doi.org/10.1002/jhm.2009)] [Medline: [23440923](https://pubmed.ncbi.nlm.nih.gov/23440923/)]
29. Barinov L, Jairaj A, Becker M, Seymour S, Lee E, Schram A, et al. Impact of Data Presentation on Physician Performance Utilizing Artificial Intelligence-Based Computer-Aided Diagnosis and Decision Support Systems. *J Digit Imaging* 2019 Jun;32(3):408-416 [FREE Full text] [doi: [10.1007/s10278-018-0132-5](https://doi.org/10.1007/s10278-018-0132-5)] [Medline: [30324429](https://pubmed.ncbi.nlm.nih.gov/30324429/)]
30. Beaudoin M, Kabanza F, Nault V, Valiquette L. Evaluation of a machine learning capability for a clinical decision support system to enhance antimicrobial stewardship programs. *Artif Intell Med* 2016 Mar;68:29-36. [doi: [10.1016/j.artmed.2016.02.001](https://doi.org/10.1016/j.artmed.2016.02.001)] [Medline: [26947174](https://pubmed.ncbi.nlm.nih.gov/26947174/)]
31. Bien N, Rajpurkar P, Ball RL, Irvin J, Park A, Jones E, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. *PLoS Med* 2018 Nov;15(11):e1002699 [FREE Full text] [doi: [10.1371/journal.pmed.1002699](https://doi.org/10.1371/journal.pmed.1002699)] [Medline: [30481176](https://pubmed.ncbi.nlm.nih.gov/30481176/)]
32. Brennan M, Puri S, Ozrazgat-Baslanti T, Feng Z, Ruppert M, Hashemighouchani H, et al. Comparing clinical judgment with the MySurgeryRisk algorithm for preoperative risk assessment: A pilot usability study. *Surgery* 2019 May;165(5):1035-1045 [FREE Full text] [doi: [10.1016/j.surg.2019.01.002](https://doi.org/10.1016/j.surg.2019.01.002)] [Medline: [30792011](https://pubmed.ncbi.nlm.nih.gov/30792011/)]
33. Chen D, Wu L, Li Y, Zhang J, Liu J, Huang L, et al. Comparing blind spots of unsedated ultrafine, sedated, and unsedated conventional gastroscopy with and without artificial intelligence: a prospective, single-blind, 3-parallel-group, randomized, single-center trial. *Gastrointest Endosc* 2020 Feb;91(2):332-339.e3. [doi: [10.1016/j.gie.2019.09.016](https://doi.org/10.1016/j.gie.2019.09.016)] [Medline: [31541626](https://pubmed.ncbi.nlm.nih.gov/31541626/)]
34. Connell A, Montgomery H, Martin P, Nightingale C, Sadeghi-Alavijeh O, King D, et al. Evaluation of a digitally-enabled care pathway for acute kidney injury management in hospital emergency admissions. *NPJ Digit Med* 2019;2:67 [FREE Full text] [doi: [10.1038/s41746-019-0100-6](https://doi.org/10.1038/s41746-019-0100-6)] [Medline: [31396561](https://pubmed.ncbi.nlm.nih.gov/31396561/)]
35. Eshel Y, Hourri-Yafin A, Benkuzari H, Lezmy N, Soni M, Charles M, et al. Evaluation of the Parasight Platform for Malaria Diagnosis. *J Clin Microbiol* 2017 Mar;55(3):768-775 [FREE Full text] [doi: [10.1128/JCM.02155-16](https://doi.org/10.1128/JCM.02155-16)] [Medline: [27974542](https://pubmed.ncbi.nlm.nih.gov/27974542/)]
36. Giannini HM, Ginestra JC, Chivers C, Draugelis M, Hanish A, Schweickert WD, et al. A Machine Learning Algorithm to Predict Severe Sepsis and Septic Shock: Development, Implementation, and Impact on Clinical Practice. *Crit Care Med* 2019 Nov;47(11):1485-1492. [doi: [10.1097/CCM.0000000000003891](https://doi.org/10.1097/CCM.0000000000003891)] [Medline: [31389839](https://pubmed.ncbi.nlm.nih.gov/31389839/)]
37. Ginestra JC, Giannini HM, Schweickert WD, Meadows L, Lynch MJ, Pavan K, et al. Clinician Perception of a Machine Learning-Based Early Warning System Designed to Predict Severe Sepsis and Septic Shock*. *Critical Care Medicine* 2019;47(11):1477-1484. [doi: [10.1097/ccm.0000000000003803](https://doi.org/10.1097/ccm.0000000000003803)]
38. Gómez-Vallejo H, Uriel-Latorre B, Sande-Meijide M, Villamarín-Bello B, Pavón R, Fdez-Riverola F, et al. A case-based reasoning system for aiding detection and classification of nosocomial infections. *Decision Support Systems* 2016 Apr;84:104-116. [doi: [10.1016/j.dss.2016.02.005](https://doi.org/10.1016/j.dss.2016.02.005)]
39. Grunwald IQ, Ragoschke-Schumm A, Kettner M, Schwindling L, Roumia S, Helwig S, et al. First Automated Stroke Imaging Evaluation via Electronic Alberta Stroke Program Early CT Score in a Mobile Stroke Unit. *Cerebrovasc Dis* 2016;42(5-6):332-338. [doi: [10.1159/000446861](https://doi.org/10.1159/000446861)] [Medline: [27304197](https://pubmed.ncbi.nlm.nih.gov/27304197/)]
40. Kanagasingam Y, Xiao D, Vignarajan J, Preetham A, Tay-Kearney M, Mehrotra A. Evaluation of Artificial Intelligence-Based Grading of Diabetic Retinopathy in Primary Care. *JAMA Netw Open* 2018 Sep 07;1(5):e182665 [FREE Full text] [doi: [10.1001/jamanetworkopen.2018.2665](https://doi.org/10.1001/jamanetworkopen.2018.2665)] [Medline: [30646178](https://pubmed.ncbi.nlm.nih.gov/30646178/)]

41. Keel S, Lee PY, Scheetz J, Li Z, Kotowicz MA, MacIsaac RJ, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Sci Rep* 2018 Mar 12;8(1):4330 [FREE Full text] [doi: [10.1038/s41598-018-22612-2](https://doi.org/10.1038/s41598-018-22612-2)] [Medline: [29531299](https://pubmed.ncbi.nlm.nih.gov/29531299/)]
42. Kiani A, Uyumazturk B, Rajpurkar P, Wang A, Gao R, Jones E, et al. Impact of a deep learning assistant on the histopathologic classification of liver cancer. *NPJ Digit Med* 2020;3:23 [FREE Full text] [doi: [10.1038/s41746-020-0232-8](https://doi.org/10.1038/s41746-020-0232-8)] [Medline: [32140566](https://pubmed.ncbi.nlm.nih.gov/32140566/)]
43. Lagani V, Chiarugi F, Manousos D, Verma V, Fursse J, Marias K, et al. Realization of a service for the long-term risk assessment of diabetes-related complications. *J Diabetes Complications* 2015 Jul;29(5):691-698. [doi: [10.1016/j.jdiacomp.2015.03.011](https://doi.org/10.1016/j.jdiacomp.2015.03.011)] [Medline: [25953402](https://pubmed.ncbi.nlm.nih.gov/25953402/)]
44. Lin H, Li R, Liu Z, Chen J, Yang Y, Chen H, et al. Diagnostic Efficacy and Therapeutic Decision-making Capacity of an Artificial Intelligence Platform for Childhood Cataracts in Eye Clinics: A Multicentre Randomized Controlled Trial. *EClinicalMedicine* 2019 Mar;9:52-59 [FREE Full text] [doi: [10.1016/j.eclinm.2019.03.001](https://doi.org/10.1016/j.eclinm.2019.03.001)] [Medline: [31143882](https://pubmed.ncbi.nlm.nih.gov/31143882/)]
45. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. *Proc Natl Acad Sci U S A* 2018 Nov 06;115(45):11591-11596 [FREE Full text] [doi: [10.1073/pnas.1806905115](https://doi.org/10.1073/pnas.1806905115)] [Medline: [30348771](https://pubmed.ncbi.nlm.nih.gov/30348771/)]
46. Liu W, Zhang Y, Bian X, Wang L, Yang Q, Zhang X, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020;26(1):13-19 [FREE Full text] [doi: [10.4103/sjg.SJG_377_19](https://doi.org/10.4103/sjg.SJG_377_19)] [Medline: [31898644](https://pubmed.ncbi.nlm.nih.gov/31898644/)]
47. Mango VL, Sun M, Wynn RT, Ha R. Should We Ignore, Follow, or Biopsy? Impact of Artificial Intelligence Decision Support on Breast Ultrasound Lesion Assessment. *AJR Am J Roentgenol* 2020 Jun;214(6):1445-1452. [doi: [10.2214/AJR.19.21872](https://doi.org/10.2214/AJR.19.21872)] [Medline: [32319794](https://pubmed.ncbi.nlm.nih.gov/32319794/)]
48. Martin CM, Vogel C, Grady D, Zarabzadeh A, Hederman L, Kellett J, et al. Implementation of complex adaptive chronic care: the Patient Journey Record system (PaJR). *J Eval Clin Pract* 2012 Dec;18(6):1226-1234. [doi: [10.1111/j.1365-2753.2012.01880.x](https://doi.org/10.1111/j.1365-2753.2012.01880.x)] [Medline: [22816797](https://pubmed.ncbi.nlm.nih.gov/22816797/)]
49. McCoy A, Das R. Reducing patient mortality, length of stay and readmissions through machine learning-based sepsis prediction in the emergency department, intensive care unit and hospital floor units. *BMJ Open Qual* 2017;6(2):e000158 [FREE Full text] [doi: [10.1136/bmjopen-2017-000158](https://doi.org/10.1136/bmjopen-2017-000158)] [Medline: [29450295](https://pubmed.ncbi.nlm.nih.gov/29450295/)]
50. McNamara DM, Goldberg SL, Latts L, Atieh Graham DM, Waintraub SE, Norden AD, et al. Differential impact of cognitive computing augmented by real world evidence on novice and expert oncologists. *Cancer Med* 2019 Nov;8(15):6578-6584 [FREE Full text] [doi: [10.1002/cam4.2548](https://doi.org/10.1002/cam4.2548)] [Medline: [31509353](https://pubmed.ncbi.nlm.nih.gov/31509353/)]
51. Mori Y, Kudo S, Misawa M, Saito Y, Ikematsu H, Hotta K, et al. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018 Sep 18;169(6):357-366. [doi: [10.7326/M18-0249](https://doi.org/10.7326/M18-0249)] [Medline: [30105375](https://pubmed.ncbi.nlm.nih.gov/30105375/)]
52. Nagaratnam K, Harston G, Flossmann E, Canavan C, Geraldine RC, Edwards C. Innovative use of artificial intelligence and digital communication in acute stroke pathway in response to COVID-19. *Future Healthc J* 2020 Jun;7(2):169-173 [FREE Full text] [doi: [10.7861/fhj.2020-0034](https://doi.org/10.7861/fhj.2020-0034)] [Medline: [32550287](https://pubmed.ncbi.nlm.nih.gov/32550287/)]
53. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S. Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening With an Offline Artificial Intelligence System on a Smartphone. *JAMA Ophthalmol* 2019 Aug 08;1182-1188 [FREE Full text] [doi: [10.1001/jamaophthalmol.2019.2923](https://doi.org/10.1001/jamaophthalmol.2019.2923)] [Medline: [31393538](https://pubmed.ncbi.nlm.nih.gov/31393538/)]
54. Nicolae A, Semple M, Lu L, Smith M, Chung H, Loblaw A, et al. Conventional vs machine learning-based treatment planning in prostate brachytherapy: Results of a Phase I randomized controlled trial. *Brachytherapy* 2020;19(4):470-476 [FREE Full text] [doi: [10.1016/j.brachy.2020.03.004](https://doi.org/10.1016/j.brachy.2020.03.004)] [Medline: [32317241](https://pubmed.ncbi.nlm.nih.gov/32317241/)]
55. Park A, Chute C, Rajpurkar P, Lou J, Ball RL, Shpanskaya K, et al. Deep Learning-Assisted Diagnosis of Cerebral Aneurysms Using the HeadXNet Model. *JAMA Netw Open* 2019 Jun 05;2(6):e195600 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.5600](https://doi.org/10.1001/jamanetworkopen.2019.5600)] [Medline: [31173130](https://pubmed.ncbi.nlm.nih.gov/31173130/)]
56. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. A lesson in implementation: A pre-post study of providers' experience with artificial intelligence-based clinical decision support. *Int J Med Inform* 2020 May;137:104072. [doi: [10.1016/j.ijmedinf.2019.104072](https://doi.org/10.1016/j.ijmedinf.2019.104072)] [Medline: [32200295](https://pubmed.ncbi.nlm.nih.gov/32200295/)]
57. Rostill H, Nilforooshan R, Morgan A, Barnaghi P, Ream E, Chrysanthaki T. Technology integrated health management for dementia. *Br J Community Nurs* 2018 Oct 02;23(10):502-508. [doi: [10.12968/bjcn.2018.23.10.502](https://doi.org/10.12968/bjcn.2018.23.10.502)] [Medline: [30290728](https://pubmed.ncbi.nlm.nih.gov/30290728/)]
58. Segal MM, Williams MS, Gropman AL, Torres AR, Forsyth R, Connolly AM, et al. Evidence-based decision support for neurological diagnosis reduces errors and unnecessary workup. *J Child Neurol* 2014 Apr;29(4):487-492. [doi: [10.1177/0883073813483365](https://doi.org/10.1177/0883073813483365)] [Medline: [23576414](https://pubmed.ncbi.nlm.nih.gov/23576414/)]
59. Segal MM, Athreya B, Son MBF, Tirosh I, Hausmann JS, Ang EYN, et al. Evidence-based decision support for pediatric rheumatology reduces diagnostic errors. *Pediatr Rheumatol Online J* 2016 Dec 13;14(1):67 [FREE Full text] [doi: [10.1186/s12969-016-0127-z](https://doi.org/10.1186/s12969-016-0127-z)] [Medline: [27964737](https://pubmed.ncbi.nlm.nih.gov/27964737/)]
60. Segal MM, Rahm AK, Hulse NC, Wood G, Williams JL, Feldman L, et al. Experience with Integrating Diagnostic Decision Support Software with Electronic Health Records: Benefits versus Risks of Information Sharing. *EGEMS (Wash DC)* 2017 Dec 06;5(1):23 [FREE Full text] [doi: [10.5334/egems.244](https://doi.org/10.5334/egems.244)] [Medline: [29930964](https://pubmed.ncbi.nlm.nih.gov/29930964/)]

61. Segal G, Segev A, Brom A, Lifshitz Y, Wasserstrum Y, Zimlichman E. Reducing drug prescription errors and adverse drug events by application of a probabilistic, machine-learning based clinical decision support system in an inpatient setting. *J Am Med Inform Assoc* 2019 Dec 01;26(12):1560-1565 [FREE Full text] [doi: [10.1093/jamia/ocz135](https://doi.org/10.1093/jamia/ocz135)] [Medline: [31390471](https://pubmed.ncbi.nlm.nih.gov/31390471/)]
62. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open Respir Res* 2017;4(1):e000234 [FREE Full text] [doi: [10.1136/bmjresp-2017-000234](https://doi.org/10.1136/bmjresp-2017-000234)] [Medline: [29435343](https://pubmed.ncbi.nlm.nih.gov/29435343/)]
63. Sim Y, Chung MJ, Kotter E, Yune S, Kim M, Do S, et al. Deep Convolutional Neural Network-based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology* 2020 Jan;294(1):199-209. [doi: [10.1148/radiol.2019182465](https://doi.org/10.1148/radiol.2019182465)] [Medline: [31714194](https://pubmed.ncbi.nlm.nih.gov/31714194/)]
64. Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, et al. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am J Surg Pathol* 2018 Dec;42(12):1636-1646 [FREE Full text] [doi: [10.1097/PAS.0000000000001151](https://doi.org/10.1097/PAS.0000000000001151)] [Medline: [30312179](https://pubmed.ncbi.nlm.nih.gov/30312179/)]
65. Su J, Li Z, Shao X, Ji C, Ji R, Zhou R, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020 Feb;91(2):415-424.e4. [doi: [10.1016/j.gie.2019.08.026](https://doi.org/10.1016/j.gie.2019.08.026)] [Medline: [31454493](https://pubmed.ncbi.nlm.nih.gov/31454493/)]
66. Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med* 2018 Sep;24(9):1337-1341. [doi: [10.1038/s41591-018-0147-y](https://doi.org/10.1038/s41591-018-0147-y)] [Medline: [30104767](https://pubmed.ncbi.nlm.nih.gov/30104767/)]
67. Vandenberghe ME, Scott MLJ, Scorer PW, Söderberg M, Balcerzak D, Barker C. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Sci Rep* 2017 Apr 05;7:45938 [FREE Full text] [doi: [10.1038/srep45938](https://doi.org/10.1038/srep45938)] [Medline: [28378829](https://pubmed.ncbi.nlm.nih.gov/28378829/)]
68. Voermans AM, Mewes JC, Broyles MR, Steuten LMG. Cost-Effectiveness Analysis of a Procalcitonin-Guided Decision Algorithm for Antibiotic Stewardship Using Real-World U.S. Hospital Data. *OMICS* 2019 Oct;23(10):508-515 [FREE Full text] [doi: [10.1089/omi.2019.0113](https://doi.org/10.1089/omi.2019.0113)] [Medline: [31509068](https://pubmed.ncbi.nlm.nih.gov/31509068/)]
69. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019 Oct;68(10):1813-1819 [FREE Full text] [doi: [10.1136/gutjnl-2018-317500](https://doi.org/10.1136/gutjnl-2018-317500)] [Medline: [30814121](https://pubmed.ncbi.nlm.nih.gov/30814121/)]
70. Wang SV, Rogers JR, Jin Y, DeiCicchi D, Dejene S, Connors JM, et al. Stepped-wedge randomised trial to evaluate population health intervention designed to increase appropriate anticoagulation in patients with atrial fibrillation. *BMJ Qual Saf* 2019 Oct;28(10):835-842 [FREE Full text] [doi: [10.1136/bmjqs-2019-009367](https://doi.org/10.1136/bmjqs-2019-009367)] [Medline: [31243156](https://pubmed.ncbi.nlm.nih.gov/31243156/)]
71. Wang P, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADE-DB trial): a double-blind randomised study. *The Lancet Gastroenterology & Hepatology* 2020 Apr;5(4):343-351. [doi: [10.1016/S2468-1253\(19\)30411-X](https://doi.org/10.1016/S2468-1253(19)30411-X)] [Medline: [31981517](https://pubmed.ncbi.nlm.nih.gov/31981517/)]
72. Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, et al. Effect of a Machine Learning-Derived Early Warning System for Intraoperative Hypotension vs Standard Care on Depth and Duration of Intraoperative Hypotension During Elective Noncardiac Surgery: The HYPE Randomized Clinical Trial. *JAMA* 2020 Mar 17;323(11):1052-1060 [FREE Full text] [doi: [10.1001/jama.2020.0592](https://doi.org/10.1001/jama.2020.0592)] [Medline: [32065827](https://pubmed.ncbi.nlm.nih.gov/32065827/)]
73. Wu X, Huang Y, Liu Z, Lai W, Long E, Zhang K, et al. Universal artificial intelligence platform for collaborative management of cataracts. *Br J Ophthalmol* 2019 Nov;103(11):1553-1560 [FREE Full text] [doi: [10.1136/bjophthalmol-2019-314729](https://doi.org/10.1136/bjophthalmol-2019-314729)] [Medline: [31481392](https://pubmed.ncbi.nlm.nih.gov/31481392/)]
74. Wu L, Zhang J, Zhou W, An P, Shen L, Liu J, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 2019 Dec;68(12):2161-2169 [FREE Full text] [doi: [10.1136/gutjnl-2018-317366](https://doi.org/10.1136/gutjnl-2018-317366)] [Medline: [30858305](https://pubmed.ncbi.nlm.nih.gov/30858305/)]
75. Yoo YJ, Ha EJ, Cho YJ, Kim HL, Han M, Kang SY. Computer-Aided Diagnosis of Thyroid Nodules via Ultrasonography: Initial Clinical Experience. *Korean J Radiol* 2018;19(4):665-672 [FREE Full text] [doi: [10.3348/kjr.2018.19.4.665](https://doi.org/10.3348/kjr.2018.19.4.665)] [Medline: [29962872](https://pubmed.ncbi.nlm.nih.gov/29962872/)]
76. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: John Wiley & Sons; 2019.
77. Auloge P, Cazzato RL, Ramamurthy N, de Marini P, Rousseau C, Garnon J, et al. Augmented reality and artificial intelligence-based navigation during percutaneous vertebroplasty: a pilot randomised clinical trial. *Eur Spine J* 2020 Jul;29(7):1580-1589. [doi: [10.1007/s00586-019-06054-6](https://doi.org/10.1007/s00586-019-06054-6)] [Medline: [31270676](https://pubmed.ncbi.nlm.nih.gov/31270676/)]
78. Liu C, Liu X, Wu F, Xie M, Feng Y, Hu C. Using Artificial Intelligence (Watson for Oncology) for Treatment Recommendations Amongst Chinese Patients with Lung Cancer: Feasibility Study. *J Med Internet Res* 2018 Sep 25;20(9):e11087 [FREE Full text] [doi: [10.2196/11087](https://doi.org/10.2196/11087)] [Medline: [30257820](https://pubmed.ncbi.nlm.nih.gov/30257820/)]
79. Embrace Healthcare Revolution: Strategically moving from volume to value. Deloitte. URL: <https://www2.deloitte.com/cn/en/pages/life-sciences-and-healthcare/articles/embrace-healthcare-revolution.html> [accessed 2021-04-05]
80. Sun Y. AI could alleviate China's doctor shortage. *MIT Technology Review*. 2018 Mar 21. URL: <https://www.technologyreview.com/2018/03/21/144544/ai-could-alleviate-chinas-doctor-shortage/> [accessed 2021-04-05]

81. China Is Building The Ultimate Technological Health Paradise. Or Is It? The Medical Futurist. 2019 Feb 19. URL: <https://medicalfuturist.com/china-digital-health/> [accessed 2021-04-05]
82. IDC Identifies China's Emerging Healthcare AI Trends in New CIO Perspective Report. IDC China. 2020 Apr 14. URL: <https://www.idc.com/getdoc.jsp?containerId=prCHC46215920> [accessed 2021-04-05]
83. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. The Lancet Oncology 2019 May;20(5):e262-e273. [doi: [10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4)] [Medline: [31044724](https://pubmed.ncbi.nlm.nih.gov/31044724/)]
84. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-Assisted Decision-making in Healthcare: The Application of an Ethics Framework for Big Data in Health and Research. Asian Bioeth Rev 2019 Sep 12;11(3):299-314 [FREE Full text] [doi: [10.1007/s41649-019-00096-0](https://doi.org/10.1007/s41649-019-00096-0)] [Medline: [33717318](https://pubmed.ncbi.nlm.nih.gov/33717318/)]
85. Longoni C, Bonezzi A, Morewedge C. Resistance to medical artificial intelligence. Journal of Consumer Research 2019;46(4):629-650. [doi: [10.1093/jcr/ucz013](https://doi.org/10.1093/jcr/ucz013)]
86. Shaffer VA, Probst CA, Merkle EC, Arkes HR, Medow MA. Why do patients derogate physicians who use a computer-based diagnostic support system? Med Decis Making 2013 Jan;33(1):108-118. [doi: [10.1177/0272989X12453501](https://doi.org/10.1177/0272989X12453501)] [Medline: [22820049](https://pubmed.ncbi.nlm.nih.gov/22820049/)]
87. Innovative solutions for digital health. Accenture. 2018. URL: <https://www.accenture.com/us-en/insights/health/healthcare-walking-ai-talk> [accessed 2021-04-05]
88. Carfagno J. IDx-DR, the First FDA-Approved AI System, is Growing Rapidly Internet. DocWire News. 2019. URL: <https://www.docwirenews.com/docwire-pick/future-of-medicine-picks/idx-dr-the-first-fda-approved-ai-system-is-growing-rapidly/> [accessed 2021-04-05]

Abbreviations

AI: artificial intelligence
DL: deep learning
FDA: Food and Drug Administration
ICU: intensive care unit
ML: machine learning
NLP: natural language processing
NPV: negative-predictive value
PPV: positive-predictive value
RCT: randomized controlled trial

Edited by T Kool; submitted 14.11.20; peer-reviewed by S Guo, E Loh; comments to author 11.01.21; revised version received 08.03.21; accepted 09.03.21; published 22.04.21

Please cite as:

Yin J, Ngiam KY, Teo HH

Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review

J Med Internet Res 2021;23(4):e25759

URL: <https://www.jmir.org/2021/4/e25759>

doi: [10.2196/25759](https://doi.org/10.2196/25759)

PMID:

©Jiamin Yin, Kee Yuan Ngiam, Hock Hai Teo. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 22.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.