

Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review



Albert T Young, Dominic Amara, Abhishek Bhattacharya, Maria L Wei



Artificial intelligence (AI) promises to change health care, with some studies showing proof of concept of a provider-level performance in various medical specialties. However, there are many barriers to implementing AI, including patient acceptance and understanding of AI. Patients' attitudes toward AI are not well understood. We systematically reviewed the literature on patient and general public attitudes toward clinical AI (either hypothetical or realised), including quantitative, qualitative, and mixed methods original research articles. We searched biomedical and computational databases from Jan 1, 2000, to Sept 28, 2020, and screened 2590 articles, 23 of which met our inclusion criteria. Studies were heterogeneous regarding the study population, study design, and the field and type of AI under study. Six (26%) studies assessed currently available or soon-to-be available AI tools, whereas 17 (74%) assessed hypothetical or broadly defined AI. The quality of the methods of these studies was mixed, with a frequent issue of selection bias. Overall, patients and the general public conveyed positive attitudes toward AI but had many reservations and preferred human supervision. We summarise our findings in six themes: AI concept, AI acceptability, AI relationship with humans, AI development and implementation, AI strengths and benefits, and AI weaknesses and risks. We suggest guidance for future studies, with the goal of supporting the safe, equitable, and patient-centred implementation of clinical AI.

Lancet Digit Health 2021;
3: e599-611

School of Medicine
(A T Young MD, D Amara MD)
and Department of
Dermatology (Prof M L Wei MD),
University of California,
San Francisco, San Francisco,
CA, USA; School of Medicine,
University of Michigan,
Ann Arbor, MI, USA
(A Bhattacharya BS);
Dermatology Service,
San Francisco Veterans Affairs
Medical Center, San Francisco,
CA, USA (Prof M L Wei)

Correspondence to:
Prof Maria L Wei, Department of
Dermatology, University of
California, San Francisco,
CA 94115, USA
maria.wei@ucsf.edu

Introduction

Artificial intelligence (AI), fuelled by advances in deep learning technology and the increasing digitisation of health-care data, shows potential for improving the diagnosis and treatment of many different medical conditions.¹ For instance, an AI-based tool that diagnoses skin lesions from photos might prompt patients to seek earlier care for melanoma,² or an AI tool that analyses electronic health record data might reduce antibiotic resistance by flagging patients being treated in hospital that were inappropriately being given broad-spectrum antibiotics.³ More broadly, AI has been shown to be able to function similarly to clinicians in medical imaging diagnosis, although few studies have been done in real-world clinical environments.⁴ Additionally, at least 64 AI-based medical devices and algorithms have been approved by the US Food & Drug Administration.⁵

Despite the rapid development of AI technology, AI has been implemented in few real-world settings because of the practical challenges of implementation⁶ and the absence of validation using metrics other than accuracy; metrics such as calibration and robustness are rarely calculated.⁷ To aid AI implementation, randomised trials for interventions involving AI should follow the recently updated Consolidated Standards of Reporting Trials-AI⁸ (updated in 2019) and Standard Protocol Items: Recommendations for Interventional Trials-AI⁹ (updated in 2020) guidelines to promote transparency and completeness. Moreover, there is a need to better understand the perspective of patients, who have the most at stake. There is an emerging body of literature on patients' attitudes toward AI, but there has been no systematic review on this topic. We systematically reviewed the literature on attitudes toward clinical AI. We summarised current knowledge and offered guidance for future studies, with the

ultimate goal of supporting the safe, equitable, and patient-centred implementation of AI in medicine.

Methods

Search strategy and inclusion criteria

In this systematic Review, we included any article that directly assessed patient or general public attitudes toward clinical AI (either hypothetical or realised) that reported original data, was published in 2000 or later up until Sept 28, 2020, was written in English, and reported enough data for thematic synthesis (ie, conference abstracts of quantitative studies presenting conclusions without numerical data were excluded). Where there were multiple reports of the same study (ie, conference abstract followed by an original research publication), we included only the original research publication. Clinical AI was defined as any software made to automate intelligent behaviour in a health-care setting for the purposes of diagnosis or treatment that might be directed towards patients, caregivers, or health-care providers, or a combination. Publications were excluded if the software of interest was simple or rule-based (ie, based on curated rule sets rather than autonomously learning from data), or both, or whose purpose was restricted to research, data collection, health education, medication management, population-level disease surveillance, or psychosocial support. For example, we considered a study¹⁰ that assessed the factors influencing patients' intention to use diabetes management apps, but we excluded this study because although the apps allowed patients to record data, receive health-related information, and communicate with providers, they did not automate intelligent behaviour such as establishing personalised medication dosing or triaging health complaints. We included studies that recruited patients, and those that sampled from the general population. We excluded

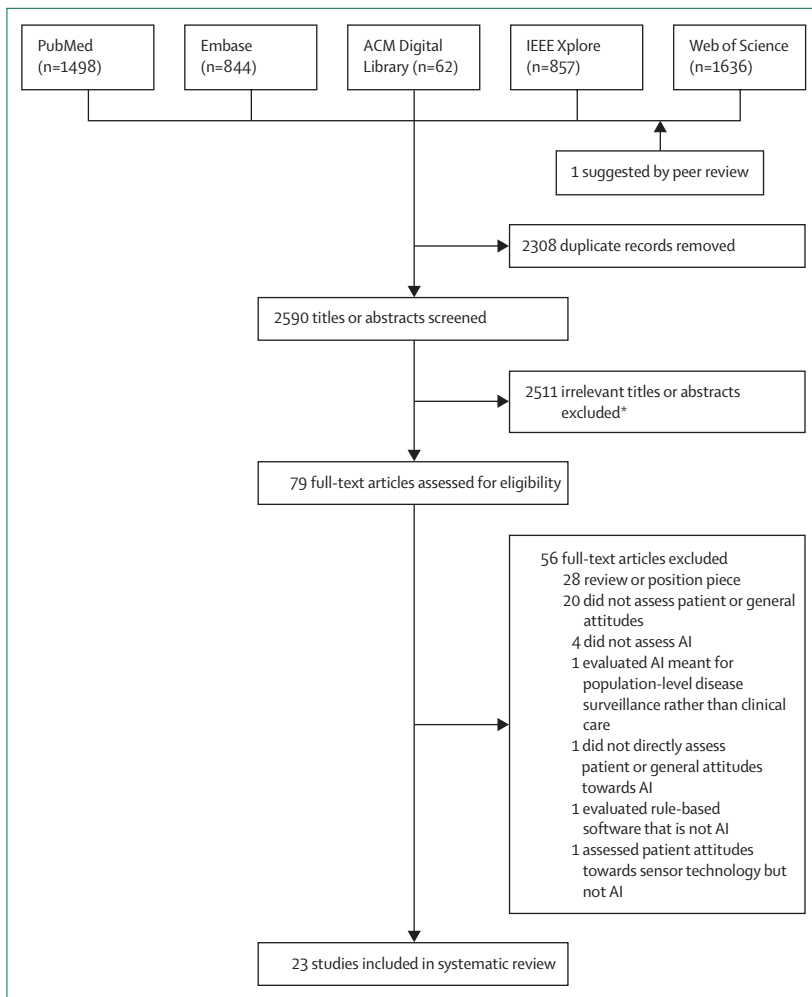


Figure: Preferred Reporting Items for Systematic reviews and Meta-Analyses flow diagram
 Reasons for the exclusion of full texts are not mutually exclusive. ACM=American for Computing Machinery, AI=artificial intelligence. IEEE=Institute of Electrical and Electronics Engineers. *Reasons for exclusion were that the title or abstract suggested the study did not assess clinical AI or did not assess patient or general attitudes, or both. The 23 studies included in the systematic review include the one found by peer review.

Science from Jan 1, 2000, to Sept 28, 2020. Search strings included terms for patients, attitudes, and AI, and are listed in the appendix (p 2). One additional article¹² missed by the search was identified and included during the peer review stage.

Data analysis

After the search, the bibliographic data were loaded into Rayyan¹³ and duplicate articles were removed. First, authors ATY and AB independently screened titles and abstracts to find out whether each could potentially meet the inclusion criteria. Articles that either reviewer decided were potentially eligible for inclusion moved on to the next stage. Second, authors ATY and DA independently assessed the full texts of the eligible articles for inclusion in the systematic Review. Third, ATY and DA together developed and piloted a data extraction form using three articles felt to represent a range of study designs among the final included articles.¹⁴⁻¹⁶ Lastly, ATY and DA independently extracted data from each included article. Disagreements at each stage were resolved by discussion with MLW. The data extraction form can be accessed online. ATY and DA critically appraised each study independently using the Mixed Methods Appraisal Tool (2018 version).¹⁷ Disagreements were resolved by discussion with MLW. We attempted to contact study authors for unclear or missing information.

We used a data-based convergent synthesis design¹⁸ to analyse all included studies using thematic synthesis. In this design, quantitative data (ie, numerical results from quantitative and mixed methods studies) are transformed into codes and analysed together with qualitative data (ie, participant quotes and themes arising from qualitative and mixed methods studies). The thematic synthesis strategy was based on a method by Thomas and Harden,¹⁹ involving free line-by-line coding, the organisation of free codes into related areas to develop descriptive themes, and finally the generation of analytical themes. Starting with a codebook based on that used by Nelson and colleagues,²⁰ ATY coded the results of each included article line by line using Dedoose.²¹ New codes were added each time a new concept was encountered, and codes were grouped into a hierarchical tree structure, leading to the development of descriptive themes. After coding all articles, ATY revisited each article to make sure codes were applied consistently. Next, DA checked this initial coding. Discrepancies were resolved through discussion with MLW and the codebook was modified accordingly. Finally, analytical themes were inferred and refined through discussion among all authors. Meta-analysis was not possible because of study heterogeneity.

**Results
Screening**

A total of 4897 records were retrieved from the electronic databases, with a total of 2590 articles left after the duplicates were removed (figure). After excluding titles

studies that sampled participants primarily based on their role as a health-care provider or health industry representative. We use the term patient to refer to participants recruited in a health-care setting and the term participant as a more general term that includes patients and those recruited in non-health-care settings.

The systematic review was registered with PROSPERO before starting data extraction (CRD42020207393). This systematic Review adheres to the Preferred Reporting Items for Systematic reviews and Meta-Analyses statement.¹¹ We first did a scoping search in Google Scholar to identify relevant articles and search terms. Under the supervision of an experienced health sciences librarian (EW, see Acknowledgements), we did a systematic literature search of PubMed, Embase, American for Computing Machinery Digital Library, Institute of Electrical and Electronics Engineers Xplore, and Web of

See Online for appendix
 For the data extraction form
 see <https://forms.gle/3QZnV7wHQG59gcqN6>

and abstracts on the basis of the inclusion criteria, 79 articles were eligible for full-text screening. A total of 23 studies fulfilled the inclusion criteria and were included in the data extraction, critical appraisal, and qualitative synthesis. Exclusion reasons are described in the appendix (pp 2–20).

	Study design	Study population, location, and response rate	Number of participants	Participant characteristics	Specialty	AI studied	Main findings
Adams et al (2020) ²²	Qualitative	Patients and family advisers and members of patient advocacy groups in Saskatchewan, Canada; response rate not reported	17	64.7% female	Radiology	AI in radiology, broadly defined	Four themes captured patients' perceptions of AI: fear of the unknown, trust, human connection, and cultural acceptability; five themes represented patient priorities for AI in radiology: improving access to imaging and reducing wait times, reducing time to diagnosis, increasing diagnostic accuracy, improving communication, and empowering patients
Bala et al (2020) ²³	Mixed methods	Patients admitted to internal medicine service and kept in hospital for at least 48 h at the University of Colorado Hospital (CO, USA), 58.8% response rate	20	55% female; 70% aged 18–34 years; 85% with at least some college education; 70% White; pregnant, incarcerated, and non-English-speaking patients were excluded	General practice	Proprietary AI software by AIPathany used to simplify medical language notes into plain language	AI-simplified medical notes were well received by patients and were more usable than un-simplified medical language notes, improved the patient-clinician relationship, and empowered patients through better understanding of their health care
Bally et al (2018) ²²	Quantitative	Patients with type 2 diabetes who required subcutaneous insulin therapy on general wards in two tertiary hospitals in the UK and Switzerland, 89% response rate	62	29% male; mean age, 67.7 years (SD 10.1)	Endocrinology	Automated fully closed-loop insulin delivery prototype (FlorenceD2W-T2)	In this randomised, open-label clinical trial, 70 patients received closed-loop insulin therapy, 62 of whom replied to a feedback questionnaire about their experience; 55/62 (89%) reported that their experience was "better than expected," and 62/62 (100%) would recommend the system to a friend or family member if they were in the hospital
Esmailzadeh (2020) ²⁶	Mixed methods	Amazon Mechanical Turk workers in the USA, response rate not reported	427	48.9% female; 30.3% aged <30 years; 90% with at least some college education; 69.7% White	General practice	AI in medicine, broadly defined	With the use of structural equation modelling, the authors found technological, ethical, and regulatory concerns to be significantly associated with the perceived risks of using clinical AI; of the factors studied, communication barriers were found to have the strongest relationship with perceived risk. Both perceived risks and benefits were significantly associated with an intention to use clinical AI
Gao et al (2020) ²⁴	Mixed methods	Users of Sina Weibo, a Chinese social media platform, in China, response rate not reported	2315 medical AI-related posts identified associated with a total of 1764 accounts, 956 of which displayed users' specific attitudes toward medical AI (number of participants not directly reported)	26.4% female; 47.5% aged <30 years; 74.6% earning more than average income	General practice	AI in medicine, broadly defined	Of 956 posts expressing attitudes toward medical AI, 568 (59.4%) expressed positive attitudes, such as the technical advantages of AI and optimism about industry development; 329 (34.4%) expressed neutral attitudes; and 59 (6.2%) expressed negative attitudes, such as concerns about the immaturity of AI technology and distrust of AI companies; of 200 posts mentioning AI replacing human doctors, 95 (47.5%) expressed that it would do so completely, 65 (32.5%) expressed that it would do so partly, and 40 (20%) expressed that it would not
Haan et al (2019) ²⁵	Qualitative	Patients scheduled for an outpatient CT scan of the chest and abdomen in the Netherlands; response rate not reported	20	Mean age 63.9 (range 39.0–79.0, SD 12.1); 45% female	Radiology	AI in radiology, broadly defined	Six domains were identified to serve as a potential framework for patient education and quantitative research: proof of technology, procedural knowledge, competence, efficiency, personal interaction, and accountability

(Table 1 continues on next page)

	Study design	Study population, location, and response rate	Number of participants	Participant characteristics	Specialty	AI studied	Main findings
(Continued from previous page)							
Juravle et al (2020) ²⁶	Quantitative	Experiment 1 included social media users and a college psychology department participant pool, and experiments 2 and 3 included users of the Testable Minds platform, in the UK, response rate not reported	379 (176 in experiment 1, 41 in experiment 2, and 162 in experiment 3)	Experiment 1: age range 18–85 years, 57.4% female; experiment 2: mean age 38.9 years (range 20–73, SD 12.86), 56.1% female; experiment 3: mean age 35.3 years (range 18–74, SD 12.29), 48.1% female	General practice	AI for medical diagnosis, broadly defined	Overall, participants trusted AI less than humans; however this gap in trust was reduced when participants were able to freely choose between the AI and human but were encouraged to choose the AI, compared with when the participant was assigned AI by authority
Jutzi et al (2020) ²⁷	Quantitative	Participants recruited through dermatological university hospital cooperation, melanoma support groups, and social media, in Germany, response rate not reported	298	75.5% aged 31–60 years; 73.2% female; 40.6% university degree; 51.7% previous diagnosis of melanoma	Dermatology	AI for medical diagnosis, first broadly defined, then specifically as an assistive tool for melanoma diagnosis	Most (94%) participants supported the use of AI in medicine in general, especially as an assistance system for physicians; only 41% supported the use of AI as a standalone system; participants with a previous history of melanoma were more amenable to using AI for the early detection of skin cancer, even at home, and they preferred an application scenario where the physician and AI classify lesions independently
Keel et al (2018) ²⁸	Quantitative	Patients attending two urban endocrinology outpatient settings, in Australia, 100% response rate on day of consultation, 57% response rate for 1 month follow-up	96	Mean age 44.26 years (range 20.00–90.00, SD 16.56); 43% female	Ophthalmology	Proprietary AI software (EyeGrader) for diagnosing diabetic retinopathy	Most (96%) participants reported that they were either satisfied or very satisfied with the automated screening model, and nearly 80% reported that they preferred the automated screening model over the manual model, suggesting a high level of acceptability to the patients
Meyer et al (2020) ²⁹	Mixed methods	Users of the Isabel symptom checker, in the USA, response rate cannot be calculated	329	Mean age 48.0 years (SD 16.7); 75.7% female; 63.2% bachelor's degree or higher; 75.3% earning less than US\$100 000 in household income; 89.1% White; 97.4% had health-care coverage; 65.7% had chronic health conditions; 59.5% had previous experience with diagnostic errors	Primary care	Isabel symptom checker	Most (90.1%) patients thought the tool provided useful information for their health problems; 51.0% reported positive health effects; the 48.4% of patients who chose to discuss the findings with their physicians conveyed mixed experiences about whether the physicians were interested in or open to discussing symptom checker results
Miller et al (2020) ³⁰	Quantitative	Patients in a large urban primary care clinic in south London, in the UK, response rate estimated to be 67%, but data not collected	523	Mean age 39.79 years (SD 17.70); 62.1% female; socioeconomic status not reported but described to have a higher than average degree of income deprivation; race not reported but clinic overall described to be 59% White	Primary care	Ada symptom checker	Most (97.8%) participants reported that Ada was very or quite easy to use, and 88.1% would use Ada again; younger participants were more likely to report that Ada had provided helpful advice, and although most (86.0%) participants reported that using Ada would not have changed their care-seeking behaviour, 12.8% reported that they would have used lower intensity care such as self-care, pharmacy, or delaying their appointment
Nadarzynski et al (2020) ³¹	Quantitative	Patients of three sexual and reproductive health clinics in Hampshire, in the UK, response rate not reported	257	73% aged 18–34 years; 57% female; 36% university degree or higher; 56% employed full time; 90% White; 25% had a previous sexually transmitted infection; 5% HIV positive; 47% had sexually transmitted infection symptoms at time of survey; 96% with internet access; 91% owned a smartphone	Sexual and reproductive health	Hypothetical AI chatbot, described as an automated webchat with a computer or a bot (not an actual human)	Less than half (40%) of participants would be willing to use an AI chatbot platform (40%) for sexual health advice, though most would be willing to use live webchats (73%) and video consultations (58%)

(Table 1 continues on next page)

	Study design	Study population, location, and response rate	Number of participants	Participant characteristics	Specialty	AI studied	Main findings
(Continued from previous page)							
Nadarzynski et al (2019) ³²	Mixed methods	Qualitative included college students responding to paper and digital adverts; quantitative included participants recruited through university-affiliated social media pages, in the UK, response rate not reported	244 (29 qualitative, 215 quantitative)	Qualitative: all aged 18–22 years, 62.5% female, 100% university students, 82.8% White; quantitative: mean age 30 years (range 18–62, SD 12), 61% female, 54% educated less than university degree, 64% White, 76% perceived to have good or very good information technology skills, 94% not aware of health chatbots	Primary care	Health chatbots, broadly defined	Most (67%) participants perceived themselves as likely to use a health chatbot within 12 months; qualitative data were organised into three themes (understanding of chatbots, AI hesitancy, and motivations for health chatbots), outlining concerns about accuracy, cybersecurity, and lack of empathy
Nelson et al (2020) ³⁰	Qualitative	Patients from general dermatology clinics in Boston, MA, USA; response rate not reported	48	Mean age 54.0 years (SD 19.9); 54% female; 77% bachelor's degree or higher; 42% had an income of \geq US\$150 000; 94% White; 94% non-Hispanic; 67% had a history of skin cancer; high percentage owned electronic devices, used digital services, and used digital services for health (eg, 81% computer ownership; 92% Google use; 81% Google use for health)	Dermatology	Direct-to-patient and clinician decision-support AI tools for skin cancer screening	Patients appeared to be receptive to the use of AI for skin cancer screening if implemented in a manner that preserves the integrity of the physician–patient relationship. Increased diagnostic speed (29 participants [60%]) and health-care access (29 participants [60%]) were the most commonly perceived benefits; increased patient anxiety was the most commonly perceived risk (19 participants [40%]); patients perceived a more accurate diagnosis (33 participants [69%]) to be the greatest strength of AI and a less accurate diagnosis (41 participants [85%]) to be the greatest weakness of AI
Ongena et al (2020) ³⁴	Quantitative	Waiting room patients scheduled for radiographic imaging, in the Netherlands, response rate not reported	155	Mean age 55.62 years (range 18.00–86.00, SD 16.56); 44.4% female; 44.8% high school education or lower	Radiology	AI in radiology that would replace doctors, broadly defined	Patients were generally not overly optimistic about AI systems taking over the diagnostic interpretations currently done by radiologists but felt that humans and AI could complement each other; patients indicated a strong need for human interaction and communication and indicated concerns about depersonalisation with the use of AI
Palmisciano et al (2020) ³³	Mixed methods	Patients who underwent brain tumour surgery, in the UK, response rate not reported	127 (20 qualitative, 107 quantitative)	Qualitative: not reported; quantitative: 52.3% \geq 46 years, 57.9% female, 55.1% completed General Certificate of Secondary Education or A level certification, 81.3% White, 59.8% religious	NeuroSurgery	AI in neurosurgery, described in five hypothetical scenarios	Most patients and their relatives thought it was acceptable to use AI for operative planning (75.8%), the real-time alert of potential complications (72.9%), and imaging interpretations (66.3%), but not partly autonomous surgery (47.7%)
Rawson et al (2019) ³⁴	Mixed methods	Attendees of a 2 day London university public festival, in the UK, response rate not reported	400 (divided into 100 groups of 2–6)	Not reported	Infectious disease	A microneedle-based biosensor and automated dose-control system for the delivery of antibiotics	Participants reported a high acceptability of the microneedle technology, but most (75%) believed that doctors should decide the antibiotic dosing, driven by concerns over computer error and the inability of AI to contextualise decision making
Spänig et al (2019) ³⁵	Quantitative	Students from three different universities, in Germany, response rate not reported	320	85% aged 24 years or less; 32.8% female; 100% university students; 37.5% agronomy, forestry, or nutritional sciences field of study; 36.6% engineering or computer science field of study	General practice	AI in medicine, broadly defined	Participants reported on average a slight positive intention to use AI for health-care needs (mean 2.6, SD 1.17, on a five-point Likert scale); a younger age and male sex were positively associated with the intention to use

(Table 1 continues on next page)

Description of included studies

Table 1 summarises the characteristics of the included studies. Of the 23 included studies, there were

14 (61%) quantitative descriptive studies, six (26%) mixed methods studies, and three (13%) qualitative studies. All studies were published between 2018 and 2020.

	Study design	Study population, location, and response rate	Number of participants	Participant characteristics	Specialty	AI studied	Main findings
(Continued from previous page)							
Stai et al (2020) ³⁶	Quantitative	Attendees of the Minnesota State Fair, in the USA, response rate not reported	264	Median age 45 years (IQR 28–59); 58% female; 70.5% bachelor's degree or higher; 72.3% with income between US\$50 000 and \$100 000; 88% White; 95.8% live in zip code where ≥70% of households have broadband internet	Urology	Hypothetical AI for assessing mass on abdominal CT, and hypothetical autonomous cancer surgery	Participants reported similar trust in the AI vs physician diagnoses; participants were generally uncomfortable with automated robotic surgery, but mistakenly believed that partly autonomous surgery was already happening; most (86%) reported that they would pay for a review of medical imaging by AI if available
Sung et al (2020) ³⁷	Quantitative	Families recruited using advertisements at a tertiary paediatric hospital, in Canada, response rate not reported	26	Not reported	Paediatrics	Hypothetical AI clinical decision support tool in low-risk (selection of antiviral therapy combination) and high-risk (invasive intervention in a patient with sepsis) scenarios	Families accepted the AI recommendation 52–92% of the time across different clinical scenarios. More families accepted the recommendation by the AI in low-risk scenarios compared with high-risk clinical scenarios
Tran et al (2019) ¹⁵	Mixed methods	Participants within the Community of Patients for Research (ComPaRe) e-cohort of patients with chronic conditions, in France, 51% response rate	1183	Raw data: median age 50 years (IQR 38–62); 73% female; 77.8% associate degree or higher; weighted data used to generate results (to reflect overall French population): median age 56 years (IQR 43–67), 54% female; 20.1% associates degree or higher	Primary care	Four existing or soon-to-be available interventions: (1) AI used to screen for skin cancer, (2) remote monitoring of chronic conditions to predict exacerbations, (3) smart clothes to guide physical therapy, (4) AI chatbots to answer emergency calls	Only 50% of patients felt that the development of digital tools and clinical AI was an important opportunity, and 11% considered it a danger. In particular, patients feared the replacement of humans and loss of the humanistic aspect of health care; 35% of patients would refuse to integrate at least one of the four interventions in their care, and few patients were ready for the use of AI without human control
Yang et al (2019) ³⁸	Quantitative	All adult inpatients from four oncology departments from two university hospitals, in China, 76.3% response rate	402	Mean age 47.86 years (14.46 SD); 50.5% female; 56.2% did not complete college; 47.8% had a family income of >5000 Chinese Yuan; 95.5% Han Chinese; 83.6% married; 95.3% religious; 65.4% reside in city; 94.6% non-medicine non-computer science occupation; 24.6% had lung cancer; 20.1% had breast cancer	Oncology	AI in medicine, broadly defined	Most participants trusted the diagnostic (90.0%) and therapeutic (85.1%) advice of AI, yet most (91.3%) would trust a human over the AI when their opinions diverge; most (87.1%) believed AI and oncologists would work together in the future, and few (11.7%) believed AI would completely replace oncologists
Ye et al (2019) ³⁹	Quantitative	Mobile telephone users, in China, 79.1% response rate	474	91.7% aged 19–40 years; 64.3% female; 83.9% bachelor's degree or higher; 39% geographically originate from Guangdong, Beijing, or Shanghai	Ophthalmology	Ophthalmic AI devices, broadly defined	With the use of structural equation modelling, the authors found subjective norms, perceived usefulness, and resistance bias to be significantly associated with the intention to use ophthalmic AI devices; the influence of subjective norms might be linked to Confucian culture, collectivism, authoritarianism, and conformity mentality in China

AI=artificial intelligence.

Table 1: Characteristics of included studies across different countries

12 (52%) studies were done in western Europe (the UK, Germany, the Netherlands, and France), seven (30%) in North America (the USA and Canada), three (13%) in China, and one (4%) in Australia. Study participants were predominantly White (59–90%) or Chinese (reported as 96% in one study). 15 studies recruited

patients in a health-care context: six from outpatient settings,^{14,20,25,28,30,31} five from inpatient settings,^{12,23,33,37,38} one from a cohort of patients with chronic conditions,¹⁵ one from users of an online symptom checker,²⁹ one from patient advocacy groups,²² and one through university hospital cooperation, melanoma support groups, and social media.²⁷ The other eight studies recruited participants outside of a health-care context: three recruited university students or affiliates, or both,^{26,32,35} and five sampled the general population.^{16,24,34,36,39} Among the quantitative and mixed methods studies, ten recruited convenience samples of participants,^{14,23,28,30,31,33–37} five did anonymous online surveys for which the response rate could not be calculated,^{16,26,27,29,32} three recruited all eligible patients,^{12,15,38} one did a simple random sampling of mobile telephone users,³⁹ and one identified all relevant social media posts.²⁴ Regarding the type of AI being studied, nine (39%) studies assessed a hypothetical AI to be used in a given clinical scenario, eight (35%) assessed AI that was broadly defined, and six (26%) assessed currently available or soon-to-be available AI tools. Regarding which specialty was represented, there were five (22%) studies for general practice, four (17%) studies for primary care, three (13%) studies for radiology, two (9%) studies for dermatology, two (9%) studies for ophthalmology, one (4%) study for endocrinology, one (4%) study for infectious disease, one (4%) study for paediatrics, one (4%) study for neurosurgery, one (4%) study for oncology, one (4%) study for sexual and reproductive health, and one (4%) study for urology.

Of the 23 studies, 21 (91%) had a section on conflicts of interests. Of those, seven (33%) studies disclosed potential or actual conflicts of interest.^{12,20,23,27,28,30,33} The quality of the methods of these studies was mixed, with all quantitative data at a high or unclear risk of selection bias, except for the single randomised trial¹² (table 2).

Thematic synthesis

The thematic synthesis yielded six broad analytical themes: (1) AI concept, (2) AI acceptability, (3) AI relationship with humans, (4) AI development and implementation, (5) AI strengths and benefits, and (6) AI weaknesses and risks. These analytical themes, which are further divided into descriptive themes, are discussed in turn. Code applications are detailed in the appendix (pp 21–41).

AI concept

Participants were generally familiar with AI^{20,27,32,33,35} but were less familiar with clinical AI.^{32,38} A few patients (3–25%) expressed having no concept of AI before the study;^{20,27,33} some responded with questioning or uncertainty.²⁵ Inpatients with cancer who were male or more highly educated reported more familiarity with AI in medicine.³⁸ Participants linked AI to the following themes: cognition (eg, game playing),²⁰

machine (eg, calculator, Google, robot, medical help telephone line),^{14,20,25,32,33} modernity,²⁰ specialised versus generalised AI,²⁰ science fiction and popular media (eg, movies and *Star Trek*),^{20,22,35} and fear.^{22,24,33}

AI acceptability

Studies measured AI acceptability in various ways, including acceptance, appropriateness, satisfaction, trust, intention to use, willingness to use again, and recommendation to family and friends. Two studies used structural equation modelling to model patients' intention to use AI based on predefined factors, such as the perceived benefits and risks.^{16,39}

Overall, participants viewed AI positively but with reservations. After using these tools in their intended clinical setting, patients expressed a high satisfaction with the EyeGrader system for detecting diabetic retinopathy (92/96 [96%] reported they were satisfied or very satisfied with automated screening, on the day of screening),²⁸ the FlorenceD2W-T2 automated fully closed-loop insulin delivery prototype (61/62 [98%] reported they were happy to have their glucose blood concentration controlled automatically),¹² and the Isabel symptom checker, which presents possible diagnoses based on symptoms entered by the user online (278/304 [91%] reported they would use it again).²⁹ Most participants supported the use of or favourably viewed AI that would be used in medicine generally,^{24,27,35,37,38} be used for symptom assessment,³⁰ be used as part of an ophthalmic device,³⁹ provide a second opinion for an imaging study,³⁶ monitor potential complications during surgery,³³ simplify medical notes,²³ screen for skin cancer,^{15,20,27} remotely monitor chronic conditions,¹⁵ guide physical therapy,¹⁵ and answer emergency telephone calls.¹⁵ However, the studies presented a diverse range of participant views; for instance, a few (approximately 20%) patients were opposed to the use of biomedical devices and AI-based tools in all four presented scenarios (the use of AI to screen for cancer, to remotely monitor chronic conditions to predict exacerbations, to guide physical therapy through smart clothes, and to answer emergency calls through chatbots),¹⁵ and 22% of patients preferred manual over automated screening for diabetic retinopathy at 1 month follow-up, citing trust as the key reason.²⁸

Participants were ambivalent toward the use of AI chatbots in health care³² and AI that would completely replace radiologists.¹⁴ Overall, the participants had negative views on automated antibiotic dosing³⁴ and were unwilling to undergo autonomous or partly autonomous surgery^{33,36} or use AI chatbots for sexual health advice.³¹ Many participants viewed currently available AI as premature technology,^{14,24,25,32,33} although some mistakenly believed that AI was already being used for surgery.³⁶

No clear themes emerged across studies about associations between AI acceptability and participant characteristics. Patients who had previously had a diagnostic error were more likely to use and report more

	Qualitative					Quantitative descriptive					Mixed methods				
	1-1	1-2	1-3	1-4	1-5	4-1	4-2	4-3	4-4	4-5	5-1	5-2	5-3	5-4	5-5
Adams et al (2020) ²²	Y	Y	Y	N	Y
Bala et al (2020) ²³	Y	Y	Y	Y	Y	C	N	Y	N	Y	Y	Y	Y	Y	N
Bally et al (2018) ¹²	Y	Y	Y	Y	Y
Esmailzadeh (2020) ¹⁵	Y	N	Y	C	Y
Gao et al (2020) ²⁴	Y	N	Y	N	Y
Haan et al (2019) ¹⁵	Y	Y	Y	N	Y
Juravle et al (2020) ²⁶	Y	C	Y	C	Y
Jutzi et al (2020) ²⁷	Y	N	Y	N	Y
Keel et al (2018) ²⁸	Y	Y	Y	N	Y
Meyer et al (2020) ²⁹	Y	Y	Y	Y	Y	Y	C	Y	C	Y	Y	Y	Y	Y	C
Miller et al (2020) ³⁰	Y	N	Y	C	Y
Nadarzynski et al (2020) ³¹	Y	Y	Y	C	Y
Nadarzynski et al (2019) ³²	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y	Y	Y	N
Nelson et al (2020) ²⁰	Y	Y	Y	Y	Y
Ongena et al (2020) ³⁴	Y	Y	Y	C	Y
Palmisciano et al (2020) ³³	Y	Y	Y	Y	Y	Y	Y	Y	C	Y	Y	Y	Y	Y	C
Rawson et al (2019) ³⁴	Y	Y	C	N	C	Y	C	Y	C	Y	Y	C	Y	Y	N
Spänig et al (2019) ³⁵	Y	N	Y	C	Y
Stai et al (2020) ³⁶	Y	C	Y	C	N
Sung et al (2020) ³⁷	C	C	Y	C	Y
Tran et al (2019) ¹⁵	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N
Yang et al (2019) ³⁸	Y	C	Y	C	Y
Ye et al (2019) ³⁹	Y	N	Y	C	C

What each number corresponds to: 1-1, is the qualitative approach appropriate to answer the research question? 1-2, are the qualitative data collection methods adequate to address the research question? 1-3, are the findings adequately derived from the data? 1-4, is the interpretation of results sufficiently substantiated by data? 1-5, is there coherence between qualitative data sources, collection, analysis, and interpretation? 4-1, is the sampling strategy relevant to address the research question? 4-2, is the sample representative of the target population? 4-3, are the measurements appropriate? 4-4, is the risk of non-response bias low? 4-5, is the statistical analysis appropriate to answer the research question? 5-1, is there an adequate rationale for using a mixed methods design to address the research question? 5-2, are the different components of the study effectively integrated to answer the research question? 5-3, are the outputs of the integration of qualitative and quantitative components adequately interpreted? 5-4, are divergencies between quantitative and qualitative results adequately addressed? 5-5, do the different components of the study adhere to the quality criteria of each tradition of the methods involved? Y=yes. N=no. C=can't tell.

Table 2: Quality evaluation of included studies using the Mixed Methods Appraisal Tool, 2018 version²⁷

health benefits from using the Isabel symptom checker.²⁹ Patients with a personal history of melanoma were more likely to support the use of AI in medicine compared with those without (97% vs 91%), but no significant differences were found by age, sex, or education.³⁷ There was a trend toward younger participants finding the Ada symptom checker to be more useful.³⁰ An intention to use AI was higher among university students who were male, younger, or studying a mathematics-related or science-related field.³⁵ A familiarity with technology was not associated with trust in AI.^{26,28}

AI acceptability was associated with many non-participant factors. Participants showed a greater acceptance of AI if they were encouraged to choose the AI over a provider,²⁶ if the AI were to be applied in a lower risk setting,^{26,32,37} if the AI was proven to be more accurate than the providers,^{25,27} if the patient's physician recommended the AI,²² or if the AI fitted societal and cultural norms.^{22,39}

AI relationship with humans

Participants strongly preferred provider supervision over AI,^{14,15,20,22,27,33,38} with few participants believing that AI would either completely replace providers^{24,27,33,38} or not be used at all.^{27,38} For example, symbiosis between providers and AI was the dominant theme in Nelson and colleagues,²⁰ and was envisioned by 94% of participants. In Yang and colleagues,³⁸ most of the patients with cancer surveyed (87.1%) believed that AI and oncologists would work together in the future, and few believed that AI would completely replace oncologists (11.7%) or not be used at all (1.2%).

The acceptability of AI generally hinged on its use as a support, rather than a replacement, of health-care providers. In Jutzi and colleagues,²⁷ 94% of participants were amenable to the use of AI as an assistance system for physicians in general, and only 41% were amenable to its use as a standalone system. In Tran and colleagues,¹⁵

80% of participants were ready to use AI in their care, but only 10–36% were ready to use AI without provider oversight. In Palmisciano and colleagues,³³ 47.7% of participants would accept partly autonomous AI surgery, but only 17.7% would accept full autonomous AI surgery; their acceptance also depended on whether the participants would receive clear and exhaustive information by the surgeon about the exact application of the AI.

Participants in many studies envisioned AI as a second opinion for providers, for example as a means to double-check providers' conclusions.^{14,20,25,26} Participants also felt that AI and providers have different advantages and could complement one another.^{14,22,27,38} For example, in Jutzi and colleagues,²⁷ nearly half the participants preferred a diagnostic routine in which the physicians and AI first classified skin lesions independently to maximise diagnostic sensitivity. In Bala and colleagues,²³ many participants felt that having both the AI-simplified medical note and the original note written by the physician would be useful.

For conflicts that arise between the AI and physician in clinical decision making, participants generally indicated that they would trust the physician.^{20,27,36,38} In the context of diagnosing skin cancer, if a discordance in opinion arose between the AI and physician, patients also commonly indicated that they would seek a biopsy,^{20,27} seek an opinion from another physician,²⁰ or seek a longitudinal follow-up from the same physician.²⁰ If the AI and providers were equally effective, patients preferred providers.^{25,27,37}

In contrast to the other studies, a study of Chinese social media users found that 47.5% of relevant posts conveyed that AI would replace all physicians, compared with 32.5% of posts conveying partial replacement only; pathologists were most frequently mentioned as those who would be replaced, followed by radiologists and dermatologists.²⁴

AI development and implementation

Participants envisioned or were asked about various specific uses of AI: diagnosis^{15,20,22,24,29} (including the detection of incidental or unrequested findings),^{14,25} treatment,^{15,24} prevention,^{15,22,24} surgery,^{33,36} delivering information,^{20,29,32} monitoring,^{15,33} recovery,^{15,24} and logistical help such as booking medical appointments.³² Patients wished to play a role in developing and implementing AI^{22,25} and to receive education about the use of AI.^{22,25,33} Regarding AI development, the participants were generally willing to share health data anonymously^{14,22,27} and expressed concern with the challenge of obtaining high-quality medical data.^{24,33}

Participants conveyed many hopes for AI implementation. These hopes included a need for AI results to be communicated clearly,^{14,22,23} the ability to ask questions,^{14,25,27} a way for AI to present results differently depending on the patient's familiarity with medical language,²³ a way for AI to continually learn when

new data become available,²² AI that is affordable,^{15,32,36} and AI that is integrated with electronic health records.²⁰ Some participants preferred AI to be set up by their physician or medical group rather than a technology company,²⁰ and some distrusted AI companies.²⁴ For skin cancer screening, patients were generally amenable to using AI at home, although less so than as a support tool for the physician.^{20,27}

The perceived challenges facing AI implementation included the need to overhaul the organisation of care,^{15,24} physician disapproval or the dismissal of patient-directed AI (35% of patients who used the Isabel symptom checker who chose not to discuss their results with their doctors agreed or strongly agreed that their doctors would disapprove of their decision to use the tool),²⁹ legal and regulatory concerns,^{16,24,25} ethical concerns,¹⁵ and liability and malpractice concerns.^{16,20} Regarding who should be liable for AI decisions, patients most often named the technology company and physician,^{20,37} and some were unsure.²⁵ Regarding who should be liable for AI data privacy, the health-care institution and technology company were most often named.²⁰

AI strengths and benefits

Participants perceived one of the primary strengths of AI to be a more accurate diagnosis,^{15,20,27,32,34,38} in part because of its ability to draw upon more data than humans.^{15,20} For instance, one patient noted that AI “has a huge database of what diagnosis A is supposed to look like as opposed to a human who only has their own life experiences”.²⁰ When used as a support tool, patients perceived that AI could boost the performance of physicians, especially those with less experience, and help physicians to learn.²⁷ Participants also valued AI for its perceived more objective diagnosis,^{15,20,27,32} more consistent diagnosis,^{20,27} more personalised care,¹⁵ more global vision of patients,¹⁵ convenience,^{15,20,27,28,32,38} and ease of use.³⁰ Another perceived strength of AI was patient activation: encouraging patients to seek health care^{20,23,29,30,32} or health information.^{20,29,32}

Participants perceived another of AI's primary benefits to be an increased efficiency because of an increased diagnostic speed,^{14,20,22,25,27,28} which is associated with lifesaving potential,²⁰ increased triage efficiency,^{15,20,22,25,27} increased labour efficiency,^{15,20,22,25,27} reduced health-care costs,^{20,24,25,29,38} fewer unnecessary procedures or therapies,^{15,20,27} and fewer unnecessary health visits.²⁹ As an example of increased diagnostic speed, one patient noted, “[t]echnology will help avoiding missing...the diagnosis of rare diseases for which the first symptoms are not always obvious. This may help doctors who are not specialists in these rare diseases”.¹⁵ Moreover, participants perceived that an increased diagnostic speed might decrease patient anxiety.^{14,15,20,22,32}

Participants perceived that AI would have the benefit of increased health-care access, a theme closely related to increased efficiency. In particular, they perceived that

AI would decrease waiting times,^{15,22,27,32} increase remote access to care,^{15,20,22} unburden the health-care system,^{20,27} increase the accessibility of data and information,¹⁵ facilitate follow-up for people with poor mobility,¹⁵ and make procedures more accessible;¹⁵ for example, one patient noted, “Care can happen everywhere. [This will help in] adjusting treatment remotely and preventing complications”.¹⁵

Patients perceived that AI could improve patient communication by increasing patient understanding of medical information,^{15,20,23,29} facilitating care when not at home with multi-language tools,¹⁵ and improving patients’ ability to communicate with non-medical people.²³ AI might empower patients by increasing autonomy,¹⁵ increasing patients’ responsibility in their care,^{15,20} and helping patients to overcome disability.²³ Participants perceived that AI could improve the patient–clinician relationship, for example by increasing the time for patient–physician interaction,^{15,20} improving follow-up and reactivity of care,^{15,23,38} and reducing patient–physician conflicts.²⁴

Other perceived potential AI benefits were increased patient privacy through anonymity,^{14,15,32} a reduced risk of medical mistakes,¹⁵ increased diagnostic transparency,²⁷ reduced health-care disparities (eg, through equitable access),^{22,24,38} the growth of AI companies²⁴ and related technologies,²⁰ and promoting health-care reform.²⁴

AI weaknesses and risks

Participants perceived one of AI’s primary weaknesses to be a less accurate diagnosis,^{14–16,20,25,32} such as the concern of missing less common health conditions.³² This perception was related to the fact that AI does not have context or human experience,^{14,20,25,34} it is unable to generalise to all individual situations,^{15,24,38} it is operator dependent,²⁰ it cannot do a physical examination,²⁰ and it has a possibly little or inaccurate training dataset.²⁰ For example, one participant expressed, “this kind of complex inspection is difficult to standardize. And it is too difficult for AI”.²⁴

Participants perceived another weakness of AI to be an inability to verbally communicate,^{20,27} to non-verbally communicate,^{20,22} and to show emotion and empathy.^{20,24,32,33} For example, patients generally agreed (mean score of 3.97 of 5 on a Likert scale) with the statement, “I find it worrisome that a computer does not take feelings into account.”¹⁴ Participants were additionally concerned with AI not having interpretability and transparency.^{14,20,27,32} For example, one patient noted, “I would probably need...feedback from a medical professional...to trust the app...because it’s like a black box...Algorithms with databases behind them...can make errors.”²⁰ Other weaknesses identified included an absence of creativity,²⁰ no social contract between the AI and the patient,²⁰ uniformity restricting the patient choice of health-care professional,²⁰ and a negative environmental effect.¹⁵

Participants perceived one of AI’s primary risks to be depersonalisation through the loss of human contact in care,^{14,15,20,24,25,27,32,38} patient loss to follow-up,²⁰ that patients are reduced to numbers,^{15,25} and the hindering of the importance given to patients’ words.¹⁵ For example, patients expressed the need for human contact to ask questions and gain mutual understanding, and they were concerned about “depersonalized procedures in which patients become numbers”.²⁵

Participants were also concerned with the risk of dependence on technology, with consequences including human deskilling^{14,20,27,24} and human job loss,^{14,24,25,32} including putting caregivers out of work,¹⁵ a reduced trust in health-care providers,²⁰ a decrease of caregiver and patient responsibility,¹⁵ and a loss of provider control in care.¹⁵ For example, patients expressed concerns that physicians would rely too much on AI such that they would lose their own diagnostic skills and not be able to recognise obvious mistakes or malfunctioning of the AI.²⁷

Furthermore, participants perceived that AI had a risk of miscommunication,^{15,16,23,32} increased patient anxiety,^{15,20,22} reduced diagnostic speed (eg, because of a delay in access to an in-person provider visit),²⁰ increased health-care costs,^{24,27,38} increased health-care disparities (through benefiting those who are most socioeconomically advantaged),^{20,27} and a risk of patient physical harm because of self-treatment¹⁵ or the use of AI technology.^{15,32} Participants were concerned about risks to privacy;^{14,15,20,27,32,33} some found privacy to be a less important concern, particularly in China.^{16,24,39} Some participants expressed concern about the inappropriate use of AI, including manipulation by hackers, data fraud, and profit-seeking,^{15,20,24,27,33} as well as technical malfunctions.^{15,27}

Discussion

In this Review, we summarise current knowledge about patient and general public attitudes toward AI using six analytical themes. The concept of AI was generally familiar through its applications outside of health care but there was less familiarity with clinical AI.

The most prevalent and consistent points identified in the literature are summarised here. Participants generally accepted the use of AI in their care but wanted proof of its effectiveness and knowledge of its exact application; a decreased interest in chatbots^{31,32} could be related to the risk of depersonalisation through AI. Participants strongly preferred to keep providers in the loop, maximising the individual strengths of health-care providers and AI. Regarding AI development and implementation, participants wished to be involved and expressed views on the potential applications and the challenges of implementation. The perceived strengths of AI included increased accuracy and efficiency, and its benefits included increased access to health care and patient empowerment. The perceived weaknesses of AI

included decreased accuracy because of an absence of appropriate algorithm training or context, an inability to communicate and empathise, and no interpretability; and the risks of AI included depersonalisation and dependence on technology.

Many themes had both positive and negative sides. Participants generally perceived accuracy to be one of the greatest strengths of AI as well as one of its greatest weaknesses, since on the one hand AI might overall improve accuracy more than that of humans alone,^{14,20,22,24,26,27,34,38} but on the other hand, its accuracy might be lower than expected in uncommon situations for which the AI is not specifically trained and for which it does not have the context or understanding that would allow it to detect mistakes.^{14,20,25,34} Participants expressed that AI could reduce health-care disparities by increasing access,^{22,24,38} but that it could worsen health-care disparities by benefiting those who are most socioeconomically advantaged.^{20,27} AI could increase privacy by allowing patients to avoid divulging sensitive mental health or sexual health information³² but decrease privacy if personal data is improperly used.^{14,15,20,33} For providers, AI could increase labour efficiency^{15,20,22,25,27} but could result in human deskilling^{14,20,24,27} and job loss.^{14,24,32} There are many more examples in which AI could lead to negative unintended consequences, and the potential benefits and harms should be carefully weighed before implementation.

A difficult challenge facing the study of patient and general public attitudes is participant recruitment, specifically for quantitative studies. All quantitative or mixed methods studies had an unclear or at least moderate risk of non-response bias, and the study populations were generally younger and better educated than the general population. The digital divide between those who are younger and older is narrowing, but still notable,⁴⁰ and older (≥ 54 years) Black and Hispanic people are less likely to use technology for health-related purposes compared with White people.⁴¹ To support the development of AI that meets the needs of a diverse population and to assess differential barriers to adoption, future studies should expand efforts to recruit participants who are older, socioeconomically disadvantaged, or from other under-represented groups. Whenever possible, patients should be recruited using consecutive sampling, the response rate should be reported, and the reasons and characteristics of those declining or unable to participate should be documented. When sampling from the general population, the random sampling of participants³⁹ might be preferred to digital snowball sampling.³² In addition, studies in developing countries are needed, because internet capacity is expanding and mobile device use is growing.^{42,43}

Our Review has some limitations. First, bias might have been introduced by inadvertently omitting studies during the search; we minimised this risk by searching several databases and using multiple synonyms to maximise the

sensitivity of our search at the same time as returning a feasible number of articles to screen. We piloted terms such as clinical decision support, computer-assisted diagnosis, computer-aided diagnosis, digital health, users, and survey, but including these in our search substantially expanded the search results without returning additional relevant articles. One article¹² was added during the peer review stage, which described a closed-loop delivery system (artificial pancreas), which was missed by our search because of the search terms used. There are other AI tools not described or classified as machine learning or AI that might have been inadvertently omitted, such as biometric monitoring devices (ie, accelerometers or pedometers, electrobiochemical sensors, or ecological momentary assessment devices)⁴⁴ that include a component of AI to act on the remote monitoring. Second, we excluded articles that were not published in English; we did not find any such articles during our scoping search. Third, using our study selection criteria, we looked only at AI used for the diagnosis or treatment of patients, or both, but we recognise there are other AI health-related tools (eg, assistive robots) that were excluded. Fourth, this systematic review found only five studies that recruited participants who actually tried using the AI being studied,^{12,23,28–30} and only three studies recruited participants who used AI in its intended clinical setting.^{12,28,29} Thus, the perspectives of participants asked about hypothetical scenarios might not fully reflect those of patients with a lived experience with AI tools. As more AI tools enter clinical practice, it will be important to assess patients' attitudes toward their real-world use and over time. It is possible that as patients' exposure to AI tools increases, their comfort with using them will also increase. It will also be important to assess health-care providers' attitudes toward AI, for which no systematic review has been done, to our knowledge.

In conclusion, we systematically reviewed patient and general public attitudes toward AI and found that participants are generally willing to accept AI in their care, but they expressed various concerns that should be addressed to successfully implement AI in clinical practice. We also identified gaps in knowledge, especially the under-representation of older and socioeconomically disadvantaged participants, to inform future research.

Contributors

ATY, AB, and MLW were responsible for the concept and design. ATY, DA, and AB did the study selection. ATY and DA did the data extraction, critical appraisal, and coding. ATY and DA drafted the manuscript. ATY, DA, AB, and MLW critically revised the manuscript for intellectually important content. All authors verified the underlying data.

Declaration of interests

We declare no competing interests.

Data sharing

The search strategy is available in the appendix (p 2). Any additional data are available upon request from the corresponding author.

Acknowledgments

We thank Evans Whitaker for his generous assistance with formulating the search and analysis strategies. This Review was funded by the

National Centre for Advancing Translational Sciences, National Institutes of Health (University of California, San Francisco, Clinical and Translational Science Institute grant number 5TL1TR001871-04 to ATY). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to the full data in the study and accept responsibility to submit for publication.

References

- 1 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44–56.
- 2 Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial intelligence in dermatology: a primer. *J Invest Dermatol* 2020; **140**: 1504–12.
- 3 Bystritsky RJ, Beltran A, Young AT, Wong A, Hu X, Doernberg SB. Machine learning for the prediction of antimicrobial stewardship intervention in hospitalized patients receiving broad-spectrum agents. *Infect Control Hosp Epidemiol* 2020; **41**: 1022–27.
- 4 Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies in medical imaging. *BMJ* 2020; **368**: m689.
- 5 Benjamins S, Dhunoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 2020; **3**: 118.
- 6 He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019; **25**: 30–36.
- 7 Young AT, Fernandez K, Pfau J, et al. Stress testing reveals gaps in clinic readiness of image-based diagnostic artificial intelligence models. *NPJ Digit Med* 2021; **4**: 10.
- 8 Liu X, Rivera SC, Faes L, et al. Reporting guidelines for clinical trials evaluating artificial intelligence interventions are needed. *Nat Med* 2019; **25**: 1467–68.
- 9 Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med* 2020; **26**: 1351–63.
- 10 Zhang Y, Liu C, Luo S, et al. Factors influencing patients' intention to use diabetes management apps based on an extended unified theory of acceptance and use of technology model: web-based survey. *J Med Internet Res* 2019; **21**: e15023.
- 11 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic reviews and Meta-Analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.
- 12 Bally L, Thabit H, Hartnell S, et al. Closed-loop insulin delivery for glycaemic control in noncritical care. *N Engl J Med* 2018; **379**: 547–56.
- 13 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016; **5**: 210.
- 14 Ongena YP, Haan M, Yakar D, Kwee TC. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol* 2020; **30**: 1033–40.
- 15 Tran V-T, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the ComPaRe e-cohort. *NPJ Digit Med* 2019; **2**: 1–8.
- 16 Esmailzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC Med Inform Decis Mak* 2020; **20**: 170.
- 17 Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf* 2018; **34**: 285–91.
- 18 Hong QN, Pluye P, Bujold M, Wassef M. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev* 2017; **6**: 61.
- 19 Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008; **8**: 45.
- 20 Nelson CA, Pérez-Chada LM, Creadore A, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. *JAMA Dermatol* 2020; **156**: 501–12.
- 21 SocioCultural Research Consultants. Dedoose version 8.3.41, web application for managing, analyzing, and presenting qualitative and mixed method research data. 2020. <https://www.dedoose.com/> (accessed Dec 29, 2020).
- 22 Adams SJ, Tang R, Babyn P. Patient perspectives and priorities regarding artificial intelligence in radiology: opportunities for patient-centered radiology. *J Am Coll Radiol* 2020; **17**: 1034–36.
- 23 Bala S, Keniston A, Burden M. Patient perception of plain-language medical notes generated using artificial intelligence software: pilot mixed-methods study. *JMIR Form Res* 2020; **4**: e16670.
- 24 Gao S, He L, Chen Y, Li D, Lai K. Public perception of artificial intelligence in medical care: content analysis of social media. *J Med Internet Res* 2020; **22**: e16649.
- 25 Haan M, Ongena YP, Hommes S, Kwee TC, Yakar D. A qualitative study to understand patient perspective on the use of artificial intelligence in radiology. *J Am Coll Radiol* 2019; **16**: 1416–19.
- 26 Juravle G, Boudouraki A, Terziyska M, Rezlescu C. Trust in artificial intelligence for medical diagnoses. *Prog Brain Res* 2020; **253**: 263–82.
- 27 Jutzi TB, Krieghoff-Henning EI, Holland-Letz T, et al. Artificial intelligence in skin cancer diagnostics: the patients' perspective. *Front Med (Lausanne)* 2020; **7**: 233.
- 28 Keel S, Lee PY, Scheetz J, 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; **8**: 4330.
- 29 Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. *J Med Internet Res* 2020; **22**: e14679.
- 30 Miller S, Gilbert S, Virani V, Wicks P. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting room: exploratory pilot study. *JMIR Human Factors* 2020; **7**: e19713.
- 31 Nadarzynski T, Bayley J, Llewellyn C, Kidsley S, Graham CA. Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice. *BMJ Sex Reprod Health* 2020; **46**: 210–17.
- 32 Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. *Digit Health* 2019; **5**: 2055207619871808.
- 33 Palmisciano P, Jamjoom AAB, Taylor D, Stoyanov D, Marcus HJ. Attitudes of patients and their relatives toward artificial intelligence in neurosurgery. *World Neurosurg* 2020; **138**: e627–33.
- 34 Rawson TM, Ming D, Gowers SA, et al. Public acceptability of computer-controlled antibiotic management: an exploration of automated dosing and opportunities for implementation. *J Infect* 2019; **78**: 75–86.
- 35 Spänig S, Emberger-Klein A, Sowa JP, Canbay A, Menrad K, Heider D. The virtual doctor: an interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes. *Artif Intell Med* 2019; **100**: 101706.
- 36 Stai B, Heller N, McSweeney S, et al. Public perceptions of artificial intelligence and robotics in medicine. *J Endourol* 2020; **34**: 1041–48.
- 37 Sung J, Portales-Casamar E, Görges M. Perceptions of expert and lay users on trust in the use of artificial intelligence for medical decision-making and risk prediction. Society for Technology in Anesthesia 2020 Annual Meeting; Austin, TX; Jan 15–18, 2020 (abstr 23).
- 38 Yang K, Zeng Z, Peng H, Jiang Y. Attitudes of Chinese cancer patients toward the clinical use of artificial intelligence. *Patient Prefer Adherence* 2019; **13**: 1867–75.
- 39 Ye T, Xue J, He M, et al. Psychosocial factors affecting artificial intelligence adoption in health care in China: cross-sectional study. *J Med Internet Res* 2019; **21**: e14316.
- 40 Anderson M, Perrin A. Technology use among seniors. May 17, 2017. <https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/> (accessed Dec 29, 2020).
- 41 Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. *Gerontologist* 2019; **59**: 6–14.

-
- 42 Purnell N. Two years ago, India lacked fast, cheap internet—one billionaire changed all that. Sept 5, 2018. https://www.wsj.com/articles/two-years-ago-india-lacked-fast-cheap-internet-one-billionaire-changed-all-that-1536159916?mod=article_inline (accessed Jan 4, 2021).
- 43 FitzGerald D. Facebook investment in Africa to expand internet capacity moves ahead. May 14, 2020. https://www.wsj.com/articles/facebook-investment-in-africa-to-expand-internet-capacity-moves-ahead-11589465142?mod=searchresults_pos19&page=2 (accessed Jan 4, 2021).
- 44 Perlmutter A, Benchoufi M, Ravaud P, Tran VT. Identification of patient perceptions that can affect the uptake of interventions using biometric monitoring devices: systematic review of randomized controlled trials. *J Med Internet Res* 2020; **22**: e18986.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.