Supplementary Table 1: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	3-4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	3-4
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	16-17
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	17
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	17
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	17-18
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	18-19
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	18-19
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	18-19
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	19
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	18-19
Synthesis	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics	17-19

Section and Topic	Item #	Checklist item	Location where item is reported
methods		and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	17-19
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	17-19
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	17-19
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	17-19
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	17-19
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	17-19
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	17-19
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	5
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	NA
Study characteristics	17	Cite each included study and present its characteristics.	Table 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Supplementary Table 3-5
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Table 2,3, Supplementary table 6-7
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Table 1, Supplementary Table 3-5
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision	NA

Section and Topic	Item #	Checklist item	Location where item is reported
		(e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	5
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	7-16
	23b	Discuss any limitations of the evidence included in the review.	7-16
	23c	Discuss any limitations of the review processes used.	14-15
	23d	Discuss implications of the results for practice, policy, and future research.	15-16
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Not registered
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Not prepared
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	NA
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	NA
Competing interests	26	Declare any competing interests of review authors.	NA
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	NA

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Supplementary Table 2: Evidence-based Mapping of Implementation Outcomes according to Proctor's Framework

	Acceptability	**Adoption**	**Appropriateness***	**Implementation Cost**	**Feasibility**	**Fidelity***	**Penetration**	**Sustainability**
Definitions	*Perception among stakeholders on the implemented DL system*	*Intention or action to employ the DL system in an evidence-based practice*	*Perceived fit or relevance of DL system in a specific clinical setting*	*Cost impact of an implementation effort*	*Extent to which the DL system can be successfully implemented in setting*	*Degree to which DL system functioned as designed*	*Integration of DL system within a service setting*	*Extent to which implementation of DL system is maintained or institutionalized within a service setting's ongoing, stable operations*
Arbabshirani, 2018 [1] **(Citation 25 in main manuscript)**		Adopted in a 24-hour outpatient clinic in USA (radiology department)	Refer to *Table S5 for main patient outcomes*		Successfully incorporated as a radiology workflow optimization tool Average processing time of the algorithm: 2.3s Significant benefit on time to diagnosis of iCH: 96% faster diagnosis in 1.5% of outpatients with unknown ICH			
Keel, 2018 [2] **(Citation 22 in main manuscript)**	Good patient acceptability - 96% of patients were satisfied - 89% of patients were likely to use the automated service again	Adopted within 2 Australian public hospital endocrinology outpatient clinics			Successfully incorporated with a fully automated retinal camera that requires minimal operator training Integrated well with routine clinical workflow: mean time per automated	Image quality for diabetic retinopathy were evaluated by the algorithm		

					screening episode ~ 7 minutes Reliable image analysis: Only 3% of participants had "ungradable" retinal images classified by the software			
Scheetz, 2021 [3] **(Citation 30 in main manuscript)**	93.7% of patients reported satisfaction and 93.2% of patients reported willingness to use AI-assisted service again	Adopted within			The AI system was functional without internet connection and performed reasonably even with the use of different camera models, suggesting its generalizability and potential in clinical real-world settings	Image quality and resolution was taken into account when using the AI-based model		
Liu, 2021 [4] **(Citation 17 in main manuscript)**		Adopted in 6 local hospitals in China (radiology department)					Deployed in 6 hospitals for COVID-19 testing during the pandemic	
Munoz-Lopez, 2021 [5] **(Citation 15 in main manuscript)**		Adopted in real-life Tele-dermatology setting in Chile				Photo qualities were graded according to "high" vs "average"		
Ruamviboonsuk, 2022 [6] **(Citation 32 in main manuscript)**	Good staff acceptability: 91.7% of nurses were satisfied with the overall experience of the DL system	Adopted in 9 primary care centres in Thailand (part of National DR screening program)			Feasible for large scale screening: successfully deployed within Thailand's national screening program Enables instant grading assessments and referral	Image quality were graded according to "ungradable" vs "gradable		

					recommendations for patients			
Liu R, 2022 [7] **(Citation 26 in main manuscript)**		Adopted in a community hospital in China			Number of referrals increased from 45 to 75 when combining fundus photography and OCT imaging	Comparisons were made to the training model and framework and technical implementation was as per designed	Deployed the dual-modality in a community hospital as a screening test	
Shah, 2022 [8] **(Citation 35 in main manuscript)**	Good patient acceptability and satisfaction: 92% of patients were satisfied and 90.4% were willing to have AI-assisted screening	Adopted in a tertiary care hospital in India			Demonstrated patient acceptability and the positive perception of AI-based screening especially given the importance of tele-screening in COVID-19	Images that did not meet the technical specifications that AI was trained in were excluded	Deployed in tertiary care setting during the pandemic	
Liu XQ, 2022 [9] **(Citation 16 in main manuscript)**		Adopted in 4 primary care stations in China			Had the ability to filter out most cases that did not require human consultation, thereby reducing human workload Demonstrated the ability to identify patients with pathology cases, allowing them to receive tele-ophthalmology advice within 1-2 days	Image flagged out if quality was "poor", requiring the patient to be rescanned	Deployed in 4 community clinics where access to ophthalmologists were lacking	
Huang, 2023 [10] **(Citation 36 in main manuscript)**	Good staff acceptability: 91% of dermatologists found SkinTeller app useful, 86% found the app efficient and accurate, 72% found	Adopted as SkinTeller app on WeChat platform			Demonstrates the potential to integrate clinical workflows and streamline processes e.g. diagnosis of lesions, assessing services,		Deployed as SkinTeller App on WeChat	

	the app helpful as a management tool for chronic disease				for dermatologists and patients with psoriasis			
Bora, 2023 [11] **(Citation 33 in main manuscript)**		Adopted in 4 centres across Thailand			Feasible for large scale deployment and was implemented at sites with high patient volumes, demonstrating more efficient allocation of resources and appropriate management.		Deployed across 5 hospitals across Thailand as part of the Thailand Prospective Study (TPS)	
Zhu, 2023 [12] **(Citation 18 in main manuscript)**		Adopted in 5 centers in New Jersey, USA			Demonstrates the potential to incorporate the DL model in community screening settings and its ability to increase efficiency and accessibility during COVID-19	Images were flagged out as "ungradable" by DL model	Deployed in 5 student-run community screening settings for underserved populations	
Antaki, 2024 [13] **(Citation 20 in main manuscript)**		Adopted in a tertiary care center in Canada		Calculated yearly savings for implementing AI in screening	Demonstrates the potential of implementing CARA as a triaging system to minimize unnecessary referrals for diabetic retinopathy	Acknowledges that current CARA model requires grader overseeing for safety reasons	Deployed with the intention of making tele-screening accessible to all patients with diabetes	
Dai, 2024 [14] **(Citation 27 in main manuscript)**		Adopted in community settings in China			Demonstratres the potential of using AI to reduce screening frequency and cost based on the severity of condition	Extensive validation was conducted across diverse settings	Deployed in environments with different ethnicities	
Gu, 2024 [15]		Adopted across 6 primary			Demonstrates the feasibility of using DL in detecting	Technical validation was	Deployed in multiple	

(Citation 28 in main manuscript)		healthcare settings in China			various retinal diseases in primary care settings	assessed along with clinical utility	primary care settings	
Li, 2024 [16] **(Citation 21 in main manuscript)**		Adopted in a primary care public institution in China			Demonstrates the potential of integrating AI to influence self-management behaviours for patients with diabetes	Technical validation was assessed along with clinical utility	Deployed in primary diabetes care setting	
Skevas, 2024 [17] **(Citation 19 in main manuscript)**		Adopted in a Ophthalmology outpatient setting at University Medical Center Hamburg-Eppendorf, Germany			Demonstrates the potential of incorporating telemedicine AI in screening for more chronic diseases than just diabetic retinopathy	Utilized human grading auditing in the workflow for AI graded results	Deployed in a setting where currently screening is only limited to diabetic retinopathy	
Wang, 2024[18] **(Citation 31 in main manuscript)**		Adopted in a Clinical Hospital Setting in China			Demonstrates the potential to incorporate DL model with a lower dose radiation CT scan in screening majority of lung conditions	Image quality was one of the key features assessed in the model	Deployed with the aim of reducing risk of radiation especially for patients who are radiation-sensitive or immunocompromised	
Chen, 2024 [19] **(Citation 29 in main manuscript)**		Adopted in EENT hospital of Fudan University			Demonstrates the potential of using the validated model in clinical management for patients with cholesteatoma			
Goessinger, 2024[20]	Good stakeholder acceptability: 95.5% of patients and 94.1% of	Adopted in the University			Demonstrates the potential to incorporate AI as an adjunct tool for melanoma			

(Citation 34 in main manuscript)	dermatologists found that AI could aid the dermatologist's diagnostic abilities	Hospital Basel in Switzerland			screening given the positive stakeholder response			
*We have used Proctor's definition for implementation outcomes but with an adaptation for how we defined appropriateness and fidelity for DL systems.								

Supplementary Table 3: Risk of Bias Analysis for Diagnostic Accuracy Studies

	Reference number in main manuscript	Consecutive/ Random sampling of patients	Case control design avoided	Appropriate Exclusion	Index test interpreted without knowledge of reference standard	Pre-specific threshold, if used	Correct classification of target condition based on reference standard	Reference standard interpreted without knowledge of index test results	Appropriate interval between index test and reference standard	Same reference standard used throughout study	All patient included in analysis
Munoz-Lopez, 2021	15	🟢	🟢	🔴	🟢	NA	🟢	🟢	🟢	🔴	🟢
Liu, 2021	17	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🟢
Scheetz , 2021	30	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🔴
Liu R, 2022	26	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🟢
Shah, 2022	35	🔴	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🔴
Liu XQ, 2022	16	🟢	🟢	NA	🟢	NA	🟢	🟢	🟢	🟢	🟢
Huang, 2023	36	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🟢
Bora, 2023	33	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🟢
Zhu, 2023	18	🟢	🟢	🟢	🟢	NA	🟢	🟢	🟢	🟢	🟢

Supplementary Table 4: Risk of Bias Analysis for Cohort Studies

	Reference number in main manuscript	Same population recruitment	Same exposure measured	Valid and reliable exposure	Confounding factors identified	Strategies to deal with confounding factors	Free of outcome at start of study	Valid and reliable outcome	Sufficient follow-up time reported	Complete follow-up	Strategies to address incomplete follow-up, if any	Statistical Analysis
Keel, 2018	22	🟢	🟢	🟠	NA	NA	🟢	🟢	🟢	🟢	🟢	🟢
Ruamviboonsuk, 2022	32	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟠	🟠	🟢
Antaki, 2024	20	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟠	🟠	🟠	🟢
Dai, 2024	27	🟢	🟢	🟢	🟠	🟠	🟢	🟢	🟢	🟠	🟠	🟢
Gu, 2024	28	🟢	🟢	🟢	🟢	🟠	🟢	🟢	🟠	🟠	🟠	🟢
Li, 2024	21	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟠	🟠	🟢
Skevas, 2024	19	🟢	🟢	🟢	🟢	🟠	🟢	🟢	🟢	🟢	🟠	🟢
Wang, 2024	31	🟢	🟢	🟢	NA	NA	🟢	🟢	🔴	🔴	NA	🟢
Chen, 2024	29	🟢	🟢	🟢	NA	NA	🟢	🟢	🟢	🟢	🟢	🟢
Goessinger, 2024	34	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🔴	🔴	NA	🟢

Supplementary Table 5: Risk of Bias Analysis for Quasi-Experimental Study

	Reference number in main manuscript	Clear cause and effect	Similar pre/post participant characteristics	Similar treatment	Control Group	Multiple outcome measurements	Complete follow-up	Comparison of outcomes measured similarly	Reliable outcome measures	Statistical Analysis
Arbabshirani, 2018	25	●	●	●	●	●	●	●	●	●

● Low risk of bias ● Unclear risk of bias ● High risk of bias

Supplementary Table 6: Service Outcome Evaluations based on IOM Standard of Care Framework

	Reference number in main manuscript	Safety	Timeliness	Efficiency	Effectiveness	Equity*	Patient-centeredness
Arbabshirani, 2018	25	✓	✓	✓	✓		
Keel, 2018	22	✓	✓	✓	✓		✓
Scheetz, 2021	30	✓		✓	✓		✓
Liu, 2021	17	✓	✓	✓	✓	✓	
Liu R, 2022	26	✓		✓	✓		
Shah, 2022	35	✓		✓	✓		✓
Munoz-Lopez, 2021	15	✓	✓		✓	✓	
Ruamviboonsuk, 2022	32	✓	✓	✓	✓	✓	
Liu XQ, 2022	16	✓	✓	✓	✓		
Huang, 2023	36	✓	✓	✓	✓		
Bora, 2023	33	✓	✓	✓	✓	✓	
Zhu, 2023	18	✓	✓	✓	✓	✓	
Antaki, 2024	20	✓		✓	✓	✓	
Dai, 2024	27		✓	✓	✓		
Gu, 2024	28	✓		✓	✓	✓	
Li, 2024	21	✓		✓	✓		✓

Skevas, 2024	19	✓		✓	✓		✓
Wang, 2024	31	✓		✓	✓	✓	
Chen, 2024	29	✓		✓	✓		
Goessinger, 2024	34	✓		✓	✓	✓	✓

* Ensuring that all individuals can access the same quality of care that the DL system provides regardless of demographics

Supplementary Table 7: Patient Outcomes in existing healthcare settings

	Clinical Outcomes					
	Accuracy	**Sensitivity**	**Specificity**	**Positive Predictive Value**	**Negative Predictive Value**	**Others**
Arbabshirani, 2018 [1] (Citation 25 in main manuscript)	Accuracy: 84% [95% CI (78%-87%)]	70% [95% CI (58%-78%)]	87% [95% CI (82%-91%)]	64%	NA	26% of cases were re-classified as emergency cases, of which 5% were identified as ICH cases Median diagnostic time of re-classified cases vs. routine cases: 19 min vs. 512 min (p<0.0001)
Keel, 2018 [2] (Citation 22 in main manuscript)	NA	92.3%	93.7%	NA	NA	Average time per automated screening visit: 6.9 minutes Majority of patients were satisfied (96%) and likely to use the DL screening service again (89%)
Scheetz, 2021 [3] (Citation 30 in main manuscript)	AUC: 0.92	96.9% [95% CI (83.8–99.9)]	87.7% [95% CI (81.8–92.2)]	59.6% [95% CI (45.1 – 73.0)]	99.3% [95% CI (96.4 - 100)]	Majority of patients (>93%) were satisfied with the AI screening model Majority of staff members reported easy usability and interpretability of AI software
Liu, 2021 [4] (Citation 17 in main manuscript)	NA	90.52%	88.51%	NA	NA	NA
Munoz-Lopez, 2021[5] (Citation 15 in main manuscript)	Based on all skin conditions presented in the cohort compared to reference standard: CNN: 41.2% Dermatologist:60.1% (p = <0.001) Residents: 57.8% (p = <0.001) General Practitioners: 49.3% (p = 0.034)	NA	NA	NA	NA	CNN deemed useful to dermatologists in 11.8% of visits CNN model impacted final diagnostic decision of dermatologist in 0.9% of cases

	Based on diagnoses that CNN was trained on compared to reference standard: CNN: 35.1% Dermatologist: 45.2% (p = 0.007) Residents: 42.6% (p = 0.045) General Practitioners: 33.1% (p = 0.582)					
Ruamviboonsuk, 2022[6] (Citation 32 in main manuscript)	DL System: 94.7% [95% CI (93%-96.2%)] Retina Specialists: 93.5% [95% CI (91.7%-95%)] (p= 0.17)	DL System: 91.4% [95% CI (87.1% - 95%%)] Retina Specialists: 84.8% [95% CI (79.4%-90%)] (p= 0.024)	DL System: 95.4% [95% CI (94.1%-96.7%)] Retina Specialists: 95.5% [95% CI (94.1%-96.7%)] (p= 0.98)	DL System: 79.2 [95% CI (73.8-84.30)] Retina Specialists: 75.6 [95% CI (69.8-81.1)]	DL System: 95.5 [95% CI (92.8-97.9)] Retina Specialists: 92.4 [95% CI (89.3-95.5)]	Percentage of nurses (n=12) who were somewhat satisfied/very satisfied with the integration of DL system in clinical workflow: Overall experience: 91.7% Convenience of screening based on time spent: 66.7% Ease of screening: 83.3% Confidence in accuracy: 75%
Liu R, 2022[7] (Citation 26 in main manuscript)	DR: 0.944 [95% CI (0.922–0.962)] DME: 0.981 [(95% CI (0.966–0.990))	DR: 91.67% [95% CI (77.5–98.2)] DME: 97.78% [95% CI (88.2–99.9)]	DR: 96.92% [95% CI (95.0–98.2)] DME: 98.38% [95% CI (96.9–99.3)]	DR: 82.98% [95% CI (73.5–89.7)] DME: 83.02% [95% CI (69.7–91.5)]	DR: 99.41% [95% CI (98.1–99.8)] DME: 99.82% [95% CI (98.8–99.9)]	Referrals increased for both ophthalmologists and AI-based fundus photography assessments when OCT images were taken into account
Shah, 2022[8] (Citation 35 in main manuscript)	NA	86.3%	96.3%	96.5%	89.4%	Majority of patients (90.4) were willing to participate in AI-based screening and acknowledged its various benefits (97.1% - 99.0%) However, 62.5% of patients felt that relying solely on AI for screening was still not feasible

Liu XQ, 2022[9] (Citation 16 in main manuscript)	Accuracy: 95.5% κ score: 0.907	For urgent: 96.6% [95% CI (91.8%-98.7%)] For both urgent and routine: 98.5% [(95% CI, 96.5%–99.4%)]	For urgent: 98.8% [95% CI (98.0%-99.3%)] For both urgent and routine: 96.2% [(95% CI, 94.6%–97.3%)]	For urgent: 91.6% [95% CI (85.7%-95.2%)] For both urgent and routine: 92.2% [(95% CI, 89.1%–94.5%)]	For urgent: 99.5% [(95% CI, 98.9%–99.8%)] For both urgent and routine: 99.3% [(95% CI, 98.4%–99.7%)]	Filtered out cases that did not require human consultation: 96.2% Time taken for patients with suspected retinal disease to receive tele-ophthalmology advice: 21.4 hours
Huang, 2023[10] (Citation 36 in main manuscript)	Mean Average Error: Dermatologists: 4.67 (2.51-7.49) Proposed DL system: 3.12	NA	NA	NA	NA	Percentage of dermatologists who were satisfied with the SkinTeller App on WeChat: Helpful/Ease of screening: 91% Confidence in efficiency and accuracy: 86% Ease of management for chronic disease: 72% Willing to recommend: 98%
Bora, 2023[11] (Citation 33 in main manuscript)	NA	90.4%	NA	NA	NA	NA
Zhu, 2023[12] (Citation 18 in main manuscript)	NA	63.2%	94.5%	32.0%	98.4%	Processing time for images: DL model: 35.6s Tele-ophthalmology grader: 129s Clinical ophthalmologist: 68s Percentage of images marked as ungradable: DL model: 38.9% Human graders: 2.6%

Antaki, 2024[13] (Citation 20 in main manuscript)	NA	For detecting referrals: 87.5% [95% CI (71.9%-95.0%)]	For detecting referrals: 66.2% [95% CI (54.3%-76.3%)]	NA	NA	Based on 5000 patients followed-up annually at the diabetes clinic, there is an estimated savings of USD $35.44 per patient if AI is implemented in the triaging system
Dai, 2024[14] (Citation 27 in main manuscript)	NA	NA	NA	NA	NA	Mean screening interval may extend from 12 months to ~32 months
Gu, 2024[15] (Citation 28 in main manuscript)	Normal fundus: 0.76 Referable diseases: 0.93-1.00 range	Normal fundus: 0.57 [95% CI (0.55- 0.59)] Referable disease (excluding retinal detachment): 0.20-1.00 range	Normal fundus: 0.99 [95% CI (0.98- 0.99)] Referable disease (excluding retinal detachment): 0.92-1.00 range	Normal fundus: 0.99 [95% CI (0.98-0.99)] Referable disease (excluding retinal detachment): 0.14-0.82 range	Normal fundus: 0.65 [95% CI (0.63-0.67)] Referable disease (excluding retinal detachmen t): 0.97-1.00 range	NA
Li, 2024[16] (Citation 21 in main manuscript)	NA	NA	NA	NA	NA	For patients diagnosed with referable diabetic retinopathy who had a combination of doctor+DL intervention, they exhibited positive self-management behaviour, in which 77.8% of them were more likely to follow-up compared to those with no DL intervention (58.44%, p =0.001) >60% patients and primary care providers demonstrated satisfaction towards the quality of recommendations by doctor+DL

Skevas, 2024[17] (Citation 19 in main manuscript)	Range: 0.85-0.90	Range: 90%-100%	Range: 76%- 80%	Range: 7.5% to 37.3%	Range: 98.9%-100%	Majority of patients were satisfied with the AI-integrated screening system, expressed ease with use of the system, and satisfaction with the length of screening
Wang, 2024[18] (Citation 31 in main manuscript)	NA	NA	NA	NA	NA	Effective radiation dose (less than 50% of the standard low-dose CT scan) along with the use of DL may be helpful in lung screening without affecting the image quality of scans
Chen, 2024[19] (Citation 29 in main manuscript)	81.8%	NA	NA	NA	NA	Integration of DL model helped with clinical decision making for 90.1% of cases
Goessinger, 2024[20] (Citation 34 in main manuscript)	NA	NA	NA	NA	NA	Percentage of dermatologists who were satisfied: Improve diagnostic performance: 94.1% Confidence in AI assessment: - For benign lesions compared to their malignancy suspicion: 1.5% - Uncertain lesions with elevated malignancy scores from AI: 46.3%

						- Overall screening examination: >90% Ease of management: 55.1% Percentage of patients who were satisfied: Confidence in AI: >92% Fear reduction of developing skin cancer: >77% Preferred a combination of AI and dermatologist examination: 61.3%

Supplementary Information: Full search strategy as of 31 December 2024

Pubmed

1. Deep Learning OR deep neural networks OR artificial neural networks

2. Implementation OR translation OR evaluation OR implementation science

3. Real-time application OR real-world application OR prospective

4. Medicine OR healthcare OR diseases OR illness

5. Full text only

6. English only

7. In humans only

8. Systematic Review OR meta-analysis OR review

Final Search: #1 AND #2 AND #3 AND #4 AND #5 AND #6 AND #7 NOT #8

Scopus

TITLE-ABS-KEY ("Deep learning" OR "deep neural networks" OR "artificial neural networks") AND TITLE-ABS-KEY ("implementation" OR "implementation science" OR "evaluation" OR "translation") AND TITLE-ABS-KEY ("real-time application" OR "real-world application") AND TITLE-ABS-KEY ("medicine" OR "healthcare" OR "diseases" OR "illness") AND NOT TITLE-ABS-KEY ("systematic review" OR "review" OR "meta-analysis") AND (LIMIT-TO (LANGUAGE , "English"))

Web of Science

TS=("deep learning" OR "deep neural networks" OR "artificial neural networks") AND TS=("medicine" OR "healthcare" OR "diseases" OR "illness") AND TS=("implementation" OR "implementation science" OR "evaluation" OR "translation") AND TS=("real-time application" OR "real-world application") NOT TS=("systematic review" OR "review" OR "meta-analysis") NOT TS=(“ANIMALS” OR “VETERINARY”)

References

1 Arbabshirani, M. R. *et al.* Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *npj Digital Medicine* **1**, 9 (2018). https://doi.org/10.1038/s41746-017-0015-z

2 Keel, S. *et al.* Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Scientific Reports* **8**, 4330 (2018). https://doi.org/10.1038/s41598-018-22612-2

3 Scheetz, J. *et al.* Real-world artificial intelligence-based opportunistic screening for diabetic retinopathy in endocrinology and indigenous healthcare settings in Australia. *Sci Rep* **11**, 15808 (2021). https://doi.org/10.1038/s41598-021-94178-5

4 Liu, B. *et al.* Assisting scalable diagnosis automatically via CT images in the combat against COVID-19. *Sci Rep* **11**, 4145 (2021). https://doi.org/10.1038/s41598-021-83424-5

5 Muñoz-López, C. *et al.* Performance of a deep neural network in teledermatology: a single-centre prospective diagnostic study. *Journal of the European Academy of Dermatology and Venereology* **35**, 546-553 (2021). https://doi.org/https://doi.org/10.1111/jdv.16979

6 Ruamviboonsuk, P. *et al.* Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. *Lancet Digit Health* **4**, e235-e244 (2022). https://doi.org/10.1016/s2589-7500(22)00017-6

7 Liu, R. *et al.* Application of artificial intelligence-based dual-modality analysis combining fundus photography and optical coherence tomography in diabetic retinopathy screening in a community hospital. *Biomed Eng Online* **21**, 47 (2022). https://doi.org/10.1186/s12938-022-01018-2

8 Shah, P., Mishra, D., Shanmugam, M., Vighnesh, M. J. & Jayaraj, H. Acceptability of artificial intelligence-based retina screening in general population. *Indian J Ophthalmol* **70**, 1140-1144 (2022). https://doi.org/10.4103/ijo.IJO_1840_21

9 Liu, X. *et al.* Evaluation of an OCT-AI-Based Telemedicine Platform for Retinal Disease Screening and Referral in a Primary Care Setting. *Transl Vis Sci Technol* **11**, 4 (2022). https://doi.org/10.1167/tvst.11.3.4

10 Huang, K. *et al.* Artificial Intelligence-Based Psoriasis Severity Assessment: Real-world Study and Application. *J Med Internet Res* **25**, e44932 (2023). https://doi.org/10.2196/44932

11 Bora, A. *et al.* Risk Stratification for Diabetic Retinopathy Screening Order Using Deep Learning: A Multicenter Prospective Study. *Transl Vis Sci Technol* **12**, 11 (2023). https://doi.org/10.1167/tvst.12.12.11

12 Zhu, A. *et al.* Implementation of deep learning artificial intelligence in vision-threatening disease screenings for an underserved community during COVID-19. *J Telemed Telecare* **30**, 1590-1597 (2024). https://doi.org/10.1177/1357633X231158832

13 Antaki, F. *et al.* Implementation of Artificial Intelligence-Based Diabetic Retinopathy Screening in a Tertiary Care Hospital in Quebec: Prospective Validation Study. *JMIR Diabetes* **9**, e59867 (2024). https://doi.org/10.2196/59867

14 Dai, L. *et al.* A deep learning system for predicting time to progression of diabetic retinopathy. *Nat Med* **30**, 584-594 (2024). https://doi.org/10.1038/s41591-023-02702-z

15 Gu, C. *et al.* Application of artificial intelligence system for screening multiple fundus diseases in Chinese primary healthcare settings: a real-world, multicentre and cross-sectional study of 4795 cases. *Br J Ophthalmol* **108**, 424-431 (2024). https://doi.org/10.1136/bjo-2022-322940

16 Li, J. *et al.* Integrated image-based deep learning and language models for primary diabetes care. *Nat Med* **30**, 2886-2896 (2024). https://doi.org/10.1038/s41591-024-03139-8

17 Skevas, C. *et al.* Implementing and evaluating a fully functional AI-enabled model for chronic eye disease screening in a real clinical environment. *BMC Ophthalmol* **24**, 51 (2024). https://doi.org/10.1186/s12886-024-03306-y

18 Wang, J. *et al.* Value of deep learning reconstruction of chest low-dose CT for image quality improvement and lung parenchyma assessment on lung window. *Eur Radiol* **34**, 1053-1064 (2024). https://doi.org/10.1007/s00330-023-10087-3

19 Chen, B. *et al.* A 3D and Explainable Artificial Intelligence Model for Evaluation of Chronic Otitis Media Based on Temporal Bone Computed Tomography: Model Development, Validation, and Clinical Application. *J Med Internet Res* **26**, e51706 (2024). https://doi.org/10.2196/51706

20 Goessinger, E. V. *et al.* Patient and dermatologists' perspectives on augmented intelligence for melanoma screening: A prospective study. *J Eur Acad Dermatol Venereol* **38**, 2240-2249 (2024). https://doi.org/10.1111/jdv.19905