IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 25, Issue 7 Ser. 3 (July. 2026), PP 72-79
www.iosrjournals.org

Artificial Intelligence And Teledermatology In Urgent Skin Cancer Referral Pathways: A Narrative Review

Author : Dr Daniel Gravino

Abstract
Background: *Skin cancer referral volumes are placing increasing strain on dermatology services. In England, over 650,000 patients were referred via the urgent suspected skin cancer two-week-wait (2WW) pathway in 2023/24, with 120,000 exceeding the 28-day diagnostic standard. Artificial intelligence (AI) and teledermatology offer candidate solutions to this capacity deficit.*
Objectives: *To critically evaluate evidence for AI-assisted lesion classification and teledermatology within urgent skin cancer referral pathways, with emphasis on diagnostic performance, cost-effectiveness, health equity, governance, and implementation — including in centralised and resource-limited healthcare systems.*
Methods: *Narrative review of PubMed, MEDLINE, EMBASE, and grey literature (NICE, NHS England, BAD). Searches were conducted January–February 2026. Of 312 records identified, 94 full texts were assessed and 48 studies or guidance documents included.*
Results: *CNN-based AI achieves diagnostic accuracy comparable with average dermatologist performance in controlled studies (pooled AUC 0.922 for melanoma). NICE conditionally recommended DERM for autonomous deployment within the NHS 2WW pathway in May 2025 — the first such designation in dermatology. However, performance disparities for Fitzpatrick V–VI skin types, absence of large-scale prospective safety data, and uncertain cost-effectiveness remain critical limitations.*
Conclusions: *AI-integrated teledermatology may improve referral pathway efficiency, particularly in centralised healthcare systems. Prospective real-world validation across diverse populations, robust governance, and mandatory skin-tone-stratified dataset standards are required before unrestricted deployment.*
Keywords: *artificial intelligence; convolutional neural network; teledermatology; skin cancer; melanoma; two-week wait; referral pathway; NHS; health equity; deep learning*

Date of Submission: 13-07-2026 Date of Acceptance: 23-07-2026

I. Introduction

Melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC) are the most common cancers among people with fair skin, and they are placing an increasing strain on health systems around the world. Since melanoma survival is still mostly determined by early diagnosis, the effectiveness and precision of urgent referral channels have direct therapeutic implications. Over 650,000 individuals in England were sent through the 2WW urgent suspected skin cancer pathway in 2023–2024; 120,000 of these patients did not receive a diagnosis or exclusion within the 28-day requirement, indicating a severely overburdened system. (1) Only 63% of the nation's 441,000 elective dermatology sessions are finished within the allotted 18 weeks. (1)

Teledermatology and artificial intelligence, particularly deep learning using convolutional neural networks (CNNs), have garnered a lot of interest as potentially complementing methods for route optimisation. Significant research has shown that, in controlled settings, AI classifiers may match or even surpass the performance of dermatologists in identifying melanoma. (2, 3) Simultaneously, the NHS and European referral routes have operationalised store-and-forward (SAF) teledermatology, proving its viability and patient acceptability. (4)

Unlike previous reviews focusing primarily on the isolated diagnostic accuracy of AI algorithms, this review critically evaluates the integration of AI within structured urgent skin cancer referral pathways — with particular focus on teledermatology implementation models, real-world performance, cost-effectiveness, health inequalities, and medico-legal governance. We additionally consider the perspective of centralised and smaller healthcare systems, including Malta, where concentrated specialist provision and geographic constraints create a specific and underexplored context for AI-teledermatology deployment. This broader focus distinguishes the present work from existing reviews and constitutes its primary contribution.

II. Methods

A narrative literature search was performed across PubMed, MEDLINE, EMBASE, and the Cochrane Library between 1 January and 28 February 2026. Search terms combined Boolean operators with MeSH headings and free-text terms: "artificial intelligence", "deep learning", "convolutional neural network", "teledermatology", "skin cancer", "melanoma", "basal cell carcinoma", "squamous cell carcinoma", "two-week wait", "urgent referral", "diagnostic accuracy", "sensitivity", "specificity", "ROC", "cost-effectiveness", "health inequalities", and "Fitzpatrick skin type". English-language publications from January 2015 to February 2026 were eligible; landmark studies predating 2015 were included where directly relevant. Grey literature was searched via NICE, NHS England, BAD, and AAD institutional websites.

Of 312 records identified, 94 full texts were assessed for eligibility, and 48 studies or guidance documents were included. Exclusion criteria comprised non-English publications, paediatric-only studies, and reports unrelated to skin lesion assessment or referral pathway implementation. Reference lists were hand-searched for additional studies.

III. Teledermatology Models And Integration Into Urgent Referral Pathways

Service Delivery Models

Three principal teledermatology modalities have been described and evaluated in the literature. Store-and-forward (SAF) involves asynchronous transmission of clinical images and patient history to a reviewing dermatologist, generating auditable image records with reported diagnostic concordance with face-to-face consultation of 72–93%, varying with image quality and clinical complexity. (5,6) Live interactive (LI) video consultation permits real-time clinician–patient interaction , high-definition uncompressed formats approach the diagnostic concordance of SAF, whilst compressed video performs significantly less well, limiting utility for lesion morphology assessment. (5) Hybrid models combine asynchronous image capture with synchronous review for equivocal cases, offering a pragmatic balance between throughput and clinical thoroughness. (7)

Whilst comparative studies generally favour SAF for dermatological triage on grounds of image quality and logistical efficiency, the practical superiority of any model is context-dependent, varying with local infrastructure, patient demographics, and clinical workflow. The systematic comparative literature remains heterogeneous, and no single model is unequivocally superior across all settings.

Teledermatology and the Two-Week-Wait Pathway

The NHS 2WW pathway mandates specialist assessment within 14 days of urgent GP referral for suspected skin cancer. Process mapping analyses have identified several structural inefficiencies in the current pathway, including delayed image acquisition, inconsistent referral information, and suboptimal clinic allocation. (8) A minority of 2WW skin referrals prove to represent confirmed malignancy — estimated at 8–10% of referred patients — underlining the potential value of intermediate risk stratification prior to face-to-face attendance. (4)

In May 2025, NICE published its Early Value Assessment (EVA) conditionally recommending DERM (Deep Ensemble for Recognition of Malignancy; Skin Analytics) as the first AI device to receive a recommendation for autonomous deployment within the NHS urgent skin cancer pathway. (9,10) NICE noted that comparative evidence suggests DERM may identify cancerous lesions with accuracy similar to teledermatology or face-to-face dermatology assessment, with potential to reduce onward referrals by approximately 30–50%. (9) However, this is a conditional rather than unconditional recommendation, and NICE specifically stipulated that patients with Fitzpatrick V and VI skin types undergo additional dermatologist review following DERM assessment, reflecting ongoing uncertainty regarding equity of performance across skin tones. (9) Over the subsequent three-year evidence-generation period, NHS England will collect data comparing outcomes of DERM plus teledermatology versus teledermatology alone, under independent oversight. (1,10)

Accessibility, Patient Perspectives, and Smaller Healthcare Systems

Patient acceptability of teledermatology is generally high. A 2025 review of patient perspectives identified faster diagnosis and reduced travel burden as principal motivators for engagement with AI-assisted diagnostic services. (7) However, patients consistently expressed reservations about AI-only pathways devoid of clinician oversight, indicating that acceptability is contingent on retained clinical governance and that the human element of the clinical encounter remains a priority for most patients. (7)

Teledermatology holds particular promise in centralised or island health systems, where specialist dermatology is concentrated at a small number of tertiary centres. In Malta — a population of approximately 550,000 across two inhabited islands — dermatology services are provided at Mater Dei Hospital and Gozo General Hospital, with additional private provision and a network of GP and family medicine practitioners serving as the primary referral source. (11) Since April 2023, Malta has operated an electronic fast-track referral pathway for suspected skin cancer, coordinating GP referrals with dedicated dermatology triage at Mater Dei Hospital. (12,13) A SAF teledermatology layer incorporating AI-assisted pre-triage would be operationally well-suited to this existing infrastructure, potentially reducing unnecessary cross-island patient travel and alleviating finite outpatient capacity — an efficiency and equity benefit with direct relevance to the Maltese context and analogous small-island health systems globally.

IV. Artificial Intelligence In Skin Lesion Classification

CNN Architecture and Transfer Learning

CNNs are supervised deep learning networks trained to classify images by learning hierarchical feature representations from labelled data — progressing from edge and texture detection in early layers to complex lesion-level morphological patterns in deeper layers. Transfer learning, whereby a CNN pre-trained on large natural image datasets is fine-tuned on dermatological images, has become standard practice, enabling competitive performance with relatively modest domain-specific training sets. (14) Ensemble architectures, combining probabilistic outputs from multiple CNNs, demonstrate incremental performance gains and form the basis of several deployed platforms, including DERM. (15) Federated learning — enabling collaborative model training across multiple institutions without centralising patient data — offers a promising mechanism for expanding training dataset diversity whilst preserving privacy, with particular relevance for acquiring representative images from currently underrepresented populations. (16)

Comparative Diagnostic Performance

The landmark study by Esteva et al. (Nature, 2017) demonstrated that a single Inception v3 CNN, trained on 129,450 clinical images across 2,032 disease categories, achieved performance on par with 21 board-certified dermatologists for both keratinocyte carcinoma and melanoma classification under standardised experimental conditions. (2) Haenssle et al. (Annals of Oncology, 2018) compared a CNN trained on over 100,000 dermoscopic images against 58 international dermatologists at two levels of clinical context; the CNN outperformed the majority of participants at both levels of assessment, including when dermatologists received supplementary clinical metadata. (3) Brinker et al. (European Journal of Cancer, 2019) similarly reported CNN superiority over 136 of 157 dermatologists in a dermoscopic melanoma classification task. (17)

Collectively, these retrospective studies demonstrate that CNN classifiers can approach or exceed average dermatologist performance in controlled image-classification tasks. A 2024 systematic review and meta-analysis in npj Digital Medicine reported pooled AI sensitivity of 0.86 (95% CI: 0.80–0.90) and specificity of 0.94 (95% CI: 0.89–0.97) with AUC 0.922 for melanoma, versus dermatologist sensitivity of 0.86 and specificity of 0.88 — approximate equivalence with a marginal AI advantage in specificity. (18) These findings must be contextualised critically. All landmark studies employed curated, biopsy-confirmed, malignancy-enriched datasets under controlled conditions that do not reflect the considerably lower malignancy prevalence of 8–10% in real-world 2WW populations. (4) At this prevalence, even high specificity implies a substantial absolute volume of false-positive referrals requiring downstream clinical management — an under-examined but clinically important limitation of comparative accuracy studies. (19) Furthermore, at least one prospective study comparing AI to dermatologists within a real-world teledermatology setting found that dermatologists outperformed the AI, underscoring the gap between controlled experimental conditions and operational clinical performance. (16)

Dermoscopy, Histopathology, and AI

AI systems trained on dermoscopic images consistently outperform those relying on clinical photographs, reflecting the richer morphological data available at magnification. (3) Dermoscopy improves melanoma sensitivity from approximately 71% with naked-eye examination to around 90% in experienced hands, and AI provides an additional increment when applied to dermoscopic data. (3) Histopathology remains the diagnostic reference standard, yet acknowledged interobserver variability among dermatopathologists for ambiguous melanocytic lesions introduces inherent uncertainty into the ground truth used to train and validate AI systems. (14) The field has not yet adequately grappled with the epistemological implications of this circularity: sensitivity and specificity estimates are only as reliable as the labels on which they are based.

V. Summary Of Major Studies

Table 1A — Diagnostic Accuracy Studies (AI vs Clinician)

Study	Year	Journal	Design	AI System	Key Metric	Finding
Esteva et al. (2)	2017	Nature	Retrospective; CNN vs. 21 dermatologists	Inception v3 CNN	ROC-AUC	CNN on par with all dermatologist experts for melanoma and BCC classification
Haenssle et al. (3)	2018	Ann Oncol	CNN vs. 58 dermatologists; two-level assessment	Custom CNN	Sensitivity / specificity	CNN outperformed majority of 58 dermatologists; advantage sustained with added clinical metadata
Brinker et al. (17)	2019	Eur J Cancer	CNN vs. 157 dermatologists	Custom CNN	Head-to-head accuracy	CNN outperformed 136/157 dermatologists for dermoscopic melanoma

						classification
Tschandl et al. (20)	2019	Lancet Oncol	CNN vs. clinicians of varying expertise	CNN ensemble	Accuracy	AI–clinician combination superior to either alone; AI outperformed non-expert clinicians
Daneshjou et al. (21)	2022	Sci Adv	Bias/equity analysis; DDI dataset	Multiple CNNs	Sensitivity by Fitzpatrick skin type	Sensitivity 0.69 (FST I–II) vs. 0.23 (FST V–VI) for DeepDerm; partial mitigation with diverse fine-tuning
Takwoingi et al. (18)	2024	npj Digital Med	Systematic review and meta-analysis	Pooled AI systems	AUC, sensitivity, specificity	AI: AUC 0.922, sensitivity 0.86, specificity 0.94; dermatologist: sensitivity 0.86, specificity 0.88

Table 1B — Implementation, Policy, and Economic Studies

Study / Source	Year	Type	Setting	Platform	Key Finding
NHS Teledermatology Roadmap (4)	2023	Policy	NHS England	N/A	SAF teledermatology endorsed for 2WW integration; standardised image acquisition required
Edge Health / UHL pilot (22)	2024	Prospective evaluation	NHS Leicester	DERM	BCR 1.05 at pilot phase; feasibility demonstrated; approximately £107K modelled savings
NICE EVA HTE24 (9)	May 2025	Health Technology Assessment	NHS England	DERM	First AI conditionally recommended for autonomous use in NHS 2WW; FST V–VI require additional dermatologist review
NIHR HTA Report (23)	2026	Health Technology Assessment	NHS	DERM	Cost–utility model developed; long-term outcome linkage identified as critical evidence gap
Malta fast-track e-referral (12,13)	2023	Service implementation	Malta, Mater Dei Hospital	Electronic referral system	Electronic fast-track referral for suspected skin cancer implemented; potential platform for AI-SAF integration

VI. Cost-Effectiveness

Economic evidence for AI-integrated teledermatology within skin cancer referral pathways is accumulating but remains preliminary in quality and generalisability. The NICE EVA (May 2025) concluded that whilst DERM may reduce onward referrals, clinical and cost-effectiveness remain uncertain due to the absence of robust prospective comparative data, and stipulated a three-year evidence-generation period with outcome monitoring under independent NHS oversight. (1,9) The NIHR-commissioned cost–utility model (2026) represents a methodological advance in linking short-term diagnostic accuracy metrics with long-term cancer outcomes, but is constrained by the same underlying data gaps. (23)

In a prospective NHS diagnostic accuracy study, projected cost savings of approximately £107,000 per trust were estimated from avoided teledermatology reviews, face-to-face assessments, and biopsies. (24) These figures represent modelled projections rather than realised savings and do not fully incorporate implementation costs including digital infrastructure, staff training, and quality assurance overheads.

The University Hospitals of Leicester pilot reported a benefit-to-cost ratio of 1.05 — near cost-neutrality at pilot phase. (22) This should not be interpreted as evidence of marginal value; benefit-to-cost ratios in early-phase technology implementation are reliably depressed by front-loaded costs and underutilised initial capacity, and neither patient-level long-term outcomes nor downstream biopsy savings were fully captured in the analysis. Comprehensive health economic modelling at realistic malignancy prevalence, incorporating full downstream pathway costs and patient-level outcomes, remains an unmet need and an explicit prerequisite for informed commissioning decisions.

VII. Health Inequalities And Dataset Bias

Underrepresentation of patients with darker skin tones in AI training datasets represents the most significant equity concern in this field. The Diverse Dermatology Images (DDI) study by Daneshjou et al. (Science Advances, 2022) demonstrated that leading CNN classifiers including Stanford's DeepDerm exhibited a near three-fold performance disparity by skin tone — sensitivity of 0.69 for Fitzpatrick types I–II versus 0.23 for types V–VI. (21) Fine-tuning on diverse data partially attenuated this gap, confirming that training dataset composition is the principal determinant of equity. (21) Given that patients from minority ethnic groups are diagnosed with skin cancers at later and more lethal stages within current clinical practice, deploying AI systems with known performance disparities at population scale risks institutionalising existing inequalities rather than mitigating them. NICE directly addressed this concern in its May 2025 DERM guidance, stipulating mandatory additional dermatologist review for all patients with Fitzpatrick V and VI skin types assessed by DERM during the evidence-generation period. (9)

Beyond skin tone, AI-teledermatology access may be restricted by digital literacy, device ownership, internet connectivity, and language. Implementation that assumes digital access risks concentrating benefit among those already advantaged. Community-based strategies, multilingual interfaces, and proactive identification of digital exclusion are necessary to ensure that efficiency gains in referral pathways translate into equitable — rather than selective — improvements in cancer detection. (7)

VIII. Ethical And Medico-Legal Considerations

AI systems used diagnostically constitute medical devices subject to formal regulatory oversight. In the UK, AI-as-a-Medical-Device (AIaMD) requires UKCA or CE marking under MHRA regulation. DERM has achieved Class III CE marking under the European Medical Device Regulation — the highest classification available — and has received a conditional NICE recommendation for autonomous use within the NHS urgent skin cancer pathway, the first dermatology AI to receive this designation, subject to ongoing evidence generation and the Fitzpatrick V–VI qualification noted above. (9,10) The EU Artificial Intelligence Act (2024) classifies medical diagnostic AI as high-risk, imposing requirements for transparency, human oversight, auditability, and post-market surveillance that place substantial obligations on both developers and deploying institutions. (25)

Medico-legal accountability when AI contributes to clinical decisions remains unresolved in most jurisdictions. Current GMC and NHS England guidance places ultimate clinical accountability with the supervising clinician regardless of AI involvement — a position that is coherent in principle but demands that clinicians possess sufficient technical literacy to critically interrogate algorithmic outputs, a capability not currently embedded in specialty training curricula. (26) The risk of automation bias — wherein clinicians defer disproportionately to AI recommendations rather than applying independent clinical judgement — is a recognised patient safety hazard and must be mitigated through training, governance structures, and deliberate system design. Informed consent should explicitly disclose that patient images will be processed by AI systems. The inherent opacity of deep learning architectures remains a barrier to genuine algorithmic transparency; explainable AI (XAI) methods such as gradient-weighted class activation mapping (Grad-CAM) offer partial solutions but are not yet standard in deployed clinical tools.

IX. Discussion

What the Evidence Establishes and What It Does Not

The evidence that CNN-based AI can approach or marginally exceed average dermatologist performance on dermoscopic melanoma classification under controlled conditions is consistent across multiple independent studies and meta-analysis. (2,3,18) However, several methodological limitations constrain the translational inference that can be drawn for real-world pathway implementation. The overwhelming majority of comparative studies employed curated, biopsy-confirmed, malignancy-enriched datasets, introducing spectrum bias — whereby disease severity in study populations systematically exceeds that encountered in clinical practice — inflating both sensitivity and specificity estimates relative to real-world performance. (19) Dataset shift, the empirically demonstrated degradation in AI performance when the deployment population differs from the training population in skin type, imaging conditions, or disease prevalence, remains insufficiently addressed and inadequately reported in primary studies. (21)

The prevalence–performance interaction deserves particular emphasis. At a 2WW malignancy prevalence of 8–10%, an AI system with pooled sensitivity of 0.86 and specificity of 0.94 would generate approximately eight false positives per true positive detected. The resulting positive predictive value of approximately 56% means that roughly half of patients flagged as high-risk and referred for further workup would prove to have benign disease — implying that the clinical and economic consequences of AI-directed triage must be modelled at realistic prevalence rather than under controlled experimental conditions. This calculation has not been adequately incorporated into the published literature and represents a critical gap in current evidence synthesis.

The clinically meaningful question, accordingly, is not whether AI outperforms dermatologists in image-classification tasks, but whether AI-augmented referral pathways improve clinically important outcomes — cancer detection rates, time to diagnosis, and ultimately mortality — compared with existing pathways. This question has not been answered by the current evidence base and remains the central research priority for the field.

Pathway Integration: Feasibility and Safety

A significant regulatory milestone that significantly advances the industry is the NICE EVA conditional recommendation for DERM (May 2025). (9, 10) It offers a structured framework for future evidence development and recognises AI's potential to lower onwards referrals while preserving patient safety. (1) However, the NICE-mandated three-year evidence-generation period reflects recognised uncertainty and should not be taken as support for early scaling. No large-scale prospective trial has addressed the most important unresolved safety question, which is whether AI-directed release of 2WW patients raises risks of missed or delayed cancer diagnosis. Deployment outside of rigorously monitored, governance-approved programs is difficult to justify clinically or medico-legally until that data is available.

Health Equity and the Structural Imperative

Instead of being incidental, the equity challenge is systemic. When implemented at population scale to a heterogeneous patient population, the almost three-fold skin-tone performance differential described by Daneshjou et al. (21) is a clinically relevant failure rather than a minor technical flaw. Post-market surveillance must include ongoing monitoring stratified by skin tone, ethnicity, and socioeconomic level, and regulatory frameworks must mandate skin-tone-stratified performance data as a prerequisite for approval. Federated learning architectures enable cooperative model improvement across institutions without centralising sensitive patient data by providing a workable method for creating the representative training datasets that existing systems lack. (16)

The Role of Dermatologists

Dermatologists must continue to play a key role in the governance, validation, and application of AI in referral pathways because they are the clinicians who are ultimately responsible for decisions made with AI assistance and because they are the subject matter experts most qualified to assess performance in a clinical setting. Responsible adoption requires active participation in clinical validation research, institutional governance committees, regulatory procedures, and the curation of AI training datasets. In order to prepare students to critically assess algorithmic outputs and carry out governance duties in AI-assisted clinical settings, future speciality and postgraduate curricula should include technical literacy in AI as a key competency.

X. Limitations

This review carries several inherent limitations. Although the narrative style allows for broad synthesis across a variety of evidence bases, it does not statistically pool data and is prone to selection bias. There is little generalisability to healthcare systems with varied configurations, particularly those of smaller island states like Malta, because the majority of the evidence base comes from high-income country contexts, mainly the UK, USA, Germany, and Australia. The current capabilities of iteratively updated deployed systems may not be reflected in specific performance data reported in studies due to the quick speed of AI development. Because studies with lower accuracy are less likely to be published in high-impact journals, publication bias probably overestimates AI performance.

Lastly, the most important evidence gap for clinical decision-making is the lack of long-term patient outcome data in the literature, especially cancer diagnosis rates in patients who were diverted from in-person evaluation by AI triage.

XI. Future Directions

Several research priorities emerge from this review:

- Prospective, multicentre randomised controlled trials comparing AI-augmented versus standard teledermatology pathways, with cancer detection rate and time to diagnosis as co-primary endpoints
- Skin-tone-stratified validation studies across ethnically diverse NHS and European populations, with Fitzpatrick type and ethnicity as mandatory reporting variables in all future AI dermatology studies
- Real-world health economic analyses incorporating full pathway costs at operational malignancy prevalence, patient-level long-term outcomes, and downstream biopsy rates
- Federated learning implementations across NHS trusts to expand training dataset diversity whilst preserving patient privacy, with particular focus on acquiring Fitzpatrick V–VI representative images
- Standardised image acquisition protocols for 2WW teledermatology submissions, with minimum quality assurance thresholds as prerequisites for AI analysis
- Explainable AI (XAI) development to support clinician interpretation of AI outputs and facilitate genuinely meaningful informed consent
- Integration studies in smaller health systems — including Malta and analogous island settings — where Malta's existing electronic fast-track referral infrastructure (12,13) could serve as a natural implementation platform for prospective AI-SAF evaluation
- Postgraduate training frameworks to ensure dermatology trainees develop sufficient technical literacy to critically evaluate AI tools and fulfil governance responsibilities in AI-assisted clinical environments

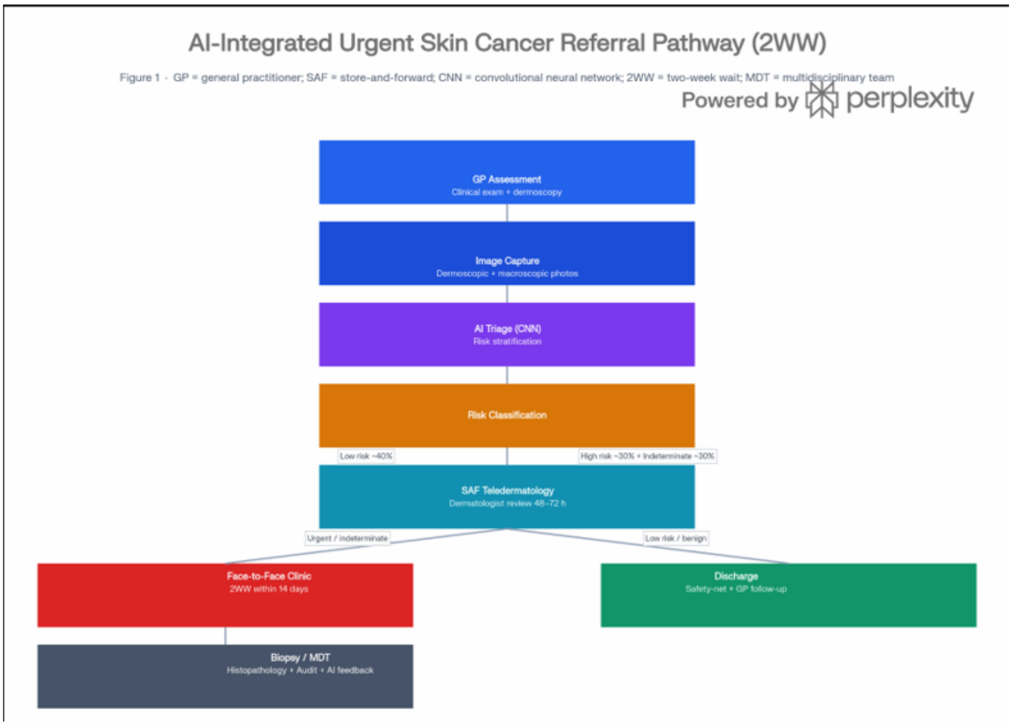

Figure 1 Legend. Proposed AI-integrated urgent skin cancer referral pathway. The model incorporates AI-assisted CNN triage at the point of GP referral, store-and-forward (SAF) teledermatology review, and risk-stratified downstream management. Approximate triage proportions at the risk classification node are illustrative, derived from NICE EVA 2025 evidence for DERM. (9) GP, general practitioner; SAF, store-and-forward teledermatology; CNN, convolutional neural network; 2WW, two-week wait; MDT, multidisciplinary team.

XII. Conclusion

The NICE conditional recommendation for DERM (May 2025) is a significant step towards operationalisation within the NHS, and AI-integrated teledermatology has the potential to increase efficiency across urgent skin cancer referral pathways. (9) Nevertheless, the information that is now available, which is primarily based on controlled, retrospective, malignancy-enriched datasets, does not yet support unrestricted autonomous deployment. Based on this review, the following clinical suggestions are put forth:

1. AI Should Function As A Supervised Triage Support Tool, Not As An Autonomous Replacement For Dermatologist Assessment, Pending Large-Scale Prospective Safety Data
2. Skin-Tone-Stratified Validation Is Required Before Population-Scale Deployment; Patients With Fitzpatrick V–VI Skin Types Require Additional Clinical Review Under Current-Generation Systems (9,21)
3. Prospective Randomised Trials Measuring Cancer Detection Rates And Time To Diagnosis Are The Priority Research Need For The Field
4. Health Economic Modelling Must Be Conducted At Operational Malignancy Prevalence, Incorporating Full Pathway Costs Rather Than Modelled Projections Under Experimental Conditions
5. Governance Frameworks Should Mandate Clinician Accountability, Algorithmic Transparency, Informed Consent, And Continuous Post-Market Outcome Monitoring Stratified By Demographic Group

AI-assisted SAF teledermatology is a good fit for the current fast-track referral infrastructure in centralised and smaller healthcare systems, such as Malta, and could enhance fair access to specialist-level triage. (12, 13) Dermatologists must continue to be involved in the governance, validation, and clinical stewardship of emerging technologies in order to appropriately realise this potential.

Declarations
Ethics approval: Not required. This is a narrative literature review and did not involve human participants.
Data availability: No new data were generated or analysed for this review.
Conflicts of interest: None declared.
Funding: None declared.
Author contributions: DG conceived the review, conducted the literature search, critically appraised the evidence, and wrote the manuscript in its entirety.

References

[1]. NHS England. AI-Based Skin Lesion Analysis Technology [Internet]. London: NHS England; 2024 [Updated December 2025; Cited February 2026]. Available From: Https://Www.England.Nhs.Uk/Elective-Care/Best-Practice-Solutions/Ai-Based-Skin-Lesion-Analysis-Technology/
[2]. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Et Al. Dermatologist-Level Classification Of Skin Cancer With Deep Neural Networks. Nature. 2017;542(7639):115–118.
[3]. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Et Al. Man Against Machine: Diagnostic Performance Of A Deep Learning Convolutional Neural Network For Dermoscopic Melanoma Recognition In Comparison To 58 Dermatologists. Ann Oncol. 2018;29(8):1836–1842.
[4]. NHS England. A Teledermatology Roadmap. London: NHS England; June 2023.
[5]. Borovicka JH, Thomas L, Todd C, Nelluri B. Teledermatology: Comparison Of Store-And-Forward Versus Live Interactive And In-Person Assessment Of Dermatological Conditions. J Telemed Telecare. 2018;24(10):649–655.
[6]. Trettel A, Eissing L, Augustin M. Telemedicine In Dermatology: Findings And Experiences Worldwide — A Systematic Literature Review. J Eur Acad Dermatol Venereol. 2018;32(2):215–224.
[7]. Hartmann D, Sun MD, Bhimireddy A, Afuwape O, Ashby I, Guo Y, Et Al. Patient Perceptions Of Artificial Intelligence And Telemedicine In Dermatology: A Narrative Review. JMIR Dermatol. 2025.
[8]. Kerr OA, Tidman MJ, Walker JJ, Aldridge RD, Benton EC. The Profile Of Dermatological Problems In Primary Care. Clin Exp Dermatol. 2010;35(4):380–383.
[9]. National Institute For Health And Care Excellence. Artificial Intelligence Technologies For Assessing And Triaging Skin Lesions Within The Urgent Suspected Skin Cancer Pathway: DERM And Moleanalyzer Pro [HTE24 Early Value Assessment]. London: NICE; May 2025.
[10]. Skin Analytics. NICE Recommends DERM For Early Use Within The NHS [Internet]. London: Skin Analytics; 2025 [Cited February 2026]. Available From: Https://Skin-Analytics.Com/News/Regulatory-Certification/Nice-Recommends-Derm-Nhs/
[11]. Maltese Association Of Dermatology And Venereology. Dermatology Services In Malta [Internet]. Msida: MADV; 2025 [Cited February 2026]. Available From: Https://Madv.Org.Mt/Patients/
[12]. Agius A. Cases Of Suspected Skin Cancer Put On Fast-Track Electronic Referrals System. Times Of Malta. 24 April 2023.
[13]. Government Of Malta — Ministry For Health. Expansion Of Fast-Track Service For Suspected Cancer Cases [Press Release]. Valletta: Department Of Information; April 2023.
[14]. Tanriverdi O, Tatliparmak A, Kaya I. Principles, Applications, And Future Of Artificial Intelligence In Dermatology. Front Med. 2023;10:1331895.
[15]. National Institute For Health And Care Excellence. Artificial Intelligence Technologies For Assessing Skin Lesions In The Urgent Suspected Cancer Pathway [HTG746]. London: NICE; 2024.
[16]. Pakzad A, Kaviani Nia A, Poortahmasebi V. AI In Dermatology: A Comprehensive Review Into Skin Cancer Detection. Dermatol Ther. 2024.
[17]. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, Et Al. Deep Learning Outperformed 136 Of 157 Dermatologists In A Head-To-Head Dermoscopic Melanoma Image Classification Task. Eur J Cancer. 2019;113:47–54.
[18]. Takwoingi Y, Chuchu N, Boachie C, Dinnes J, Matin RN, Moreau JF, Et Al. Systematic Review And Meta-Analysis Of Artificial Intelligence Versus Clinicians For Skin Cancer Diagnosis. Npj Digit Med. 2024;7:117.
[19]. Freeman K, Dinnes J, Chuchu N, Takwoingi Y, Bayliss SE, Matin RN, Et Al. Algorithm Based Smartphone Apps To Assess Risk Of Skin Cancer In Adults: Systematic Review Of Diagnostic Accuracy Studies. BMJ. 2020;368:M127.
[20]. Tschandl P, Codella N, Akay BN, Argenziano G, Braun RP, Cabo H, Et Al. Comparison Of The Accuracy Of Human Readers Versus Machine-Learning Algorithms For Pigmented Skin Lesion Classification: An Open, Web-Based, International, Diagnostic Study. Lancet Oncol. 2019;20(7):938–947.
[21]. Daneshjou R, Vodrahalli K, Novoa RA, Jenkins M, Liang W, Rotemberg V, Et Al. Disparities In Dermatology AI Performance On A Diverse, Curated Clinical Image Set. Sci Adv. 2022;8(32):Eabq6147.
[22]. Edge Health / Skin Analytics. An Evaluation Of AI-Powered Teledermatology For Skin Cancer 2WW Pathway: University Hospitals Of Leicester. London: Edge Health; 2024.
[23]. National Institute For Health Research. Artificial Intelligence Technologies For Assessing Skin Lesions For Suspected Skin Cancer [HTA Report GJMS0317]. London: NIHR; 2026.
[24]. Tam C, Polkinghorne A, Gillis R, Borys D. Accuracy Of An Artificial Intelligence As A Medical Device In The Detection Of Clinically Significant Skin Lesions In A Teledermatology Setting. Australas J Dermatol. 2024. Doi:10.1111/Ajd.14214.
[25]. European Parliament. Regulation (EU) 2024/1689 Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act). Official Journal Of The European Union; 2024.
[26]. Puri P, Comfere N, Pittelkow MR, Wieland CN, Davis MDP, Mangold AR. Deep Learning For Dermatologists: Part II. Current Applications. J Am Acad Dermatol. 2022;87(6):1352–1360.
[27]. Maron RC, Weichenthal M, Uhlmann L, Blum A, Lallas A, Argenziano G, Et Al. Systematic Outperformance Of 112 Dermatologists In Multiclass Skin Cancer Image Classification By Convolutional Neural Networks. Eur J Cancer. 2019;119:57–65.
[28]. Adamson AS, Smith A. Machine Learning And Health Care Disparities In Dermatology. JAMA Dermatol. 2018;154(11):1247–1248.
[29]. Han SS, Moon IJ, Lim W, Suh IS, Chang SE, Lee MW, Et Al. Keratinocyte Carcinoma Detection And Classification Using Artificial Intelligence. JAMA Dermatol. 2020;156(4):389–398.
[30]. Janda M, Horsham C, Vagenas D, Loescher LJ, Gillespie N, Colditz PB, Et Al. Accuracy Of Mobile Phone-Based Assessments Of Teledermatology: A Systematic Review. JAMA Dermatol. 2019;155(12):1341–1350.