Artificial intelligence (AI) technologies for assessing and triaging skin lesions referred to the urgent suspected skin cancer pathway: early value assessment

Diagnostics Consultation Document – Comments

Diagnostics Advisory Committee date: 22 October 2024

Theme: Capacity issues

Comment number	Name and organisation	Section number	Comment	NICE response
1	Consultee 1	3.1	Unmet need 3.1:- With the increasing number of cancers, especially amongst younger patients, added to the current unmet need, the wait lists will increase organically. There are three options. 1) Increase the number of patients per consultant which is impractical 2) Increase the number of Consultants which again is not possible due to their lengthy training and the demand is current and overwhelming. 3) Use AI to exclude the non malignant referrals before using Consultants time. Many of the referrals to a Dermatology Consultant are non malignant and Skin Analytics have proven AI at a 99.85% accuracy for non malignant lesion detection.	Thank you for your comment which the committee considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance. The committee also concluded that further evidence should be generated on the specificity of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone. Please see section 3.6 of the guidance for more detail.
2	Skin Analytics	1.3	While it is important to note that the use of DERM does not preclude the use of teledermatology, we are also concerned that the implicit focus on teledermatology and lack thereof on face-to-face care overlooks the workforce limitations that underpin the unmet need: - 24% of Dermatology consultant positions remained vacant, with only 508 whole-time-equivalent (WTE) consultants in posts unevenly distributed across the country and 159 WTE vacancies [8]; - Despite recommendations to increase training posts, the number fell to 32 across England in 2023 [9] - In 2023 there were >680,000 urgent suspected skin cancer referrals, which is increasing 11% year-on-year [10]	Thank you for your comment which the committee considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.

			- There is a backlog of 441,000 non urgent suspected skin cancer outpatient appointments, with only 63% seen within the 18 week target [11] - 1 in 3 melanoma and SCCs are on these non urgent suspected skin cancer pathways and these delayed diagnoses disproportionately affect older, Black and Asian patients [12]	
3	Skin Analytics	1.3	Furthermore, the statement does not reference or contextualise the known harm that already exists within cancer pathways: - 1 in 3 melanoma and SCCs are on non urgent suspected skin cancer pathways [6], where there is a backlog of 441,000 patients [7] and these delayed diagnoses disproportionately affect older, Black and Asian patients [6] 6. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection 7. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.
4	Skin Analytics	2.5	This section risks underestimating the scale of the current demand versus capacity issues and how many melanoma and SCCs are also on the non urgent suspected skin cancer pathway: - 24% of Dermatology consultant positions remained vacant, with only 508 whole-time-equivalent (WTE) consultants in posts and 159 WTE vacancies [1]; - Despite recommendations to increase training posts, the number fell to 32 across England in 2023 [2] - In 2023 there were >680,000 urgent suspected skin cancer referrals, which is increasing 11% year on year [3] - There is a backlog of 441,000 non urgent suspected skin cancer outpatient appointments, with only 63% seen within the 18 week target [4] - 1 in 3 melanoma and SCCs are on these non urgent suspected skin cancer pathways and these delayed diagnoses disproportionately affect older, Black and Asian patients [5] If not appropriately covered during the committee meetings it is possible that conclusions were reached without all of the available evidence.	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.

			1. https://gettingitrightfirsttime.co.uk/wp-content/uploads/2021/09/DermatologyReport-Sept21o.pdf 2. https://medical.hee.nhs.uk/medical-training-recruitment/medical-specialty-training/competition-ratios 3. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/&sa=D&source=editors&ust=1725441104802031&usg=AOvVaw3v5f6sjG8WDrN52MQJkFhG 4. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 5. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection	
5	Skin Analytics	3.1	This section should do more to quantify the current demand versus capacity issues and how many melanoma and SCCs are also not on the urgent suspected skin cancer pathway: - 24% of Dermatology consultant positions remained vacant, with only 508 whole-time-equivalent (WTE) consultants in posts and 159 WTE vacancies [1]; - Despite recommendations to increase training posts, the number fell to 32 across England in 2023 [2] - In 2023 there were >680,000 urgent suspected skin cancer referrals, which is increasing 11% year on year [3] - There is a backlog of 441,000 non urgent suspected skin cancer outpatient appointments, with only 63% seen within the 18 week target [4] - 1 in 3 melanoma and SCCs are on these non urgent suspected skin cancer pathways and these delayed diagnoses disproportionately affect older, Black and Asian patients [5] The omission of this data from the committee discussion may have adversely impacted their decision making when it comes to considering risks and benefits of DERM. 1. https://gettingitrightfirsttime.co.uk/wp-	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.

			content/uploads/2021/09/DermatologyReport-Sept21o.pdf 2. https://medical.hee.nhs.uk/medical-training-recruitment/medical-specialty-training/competition-ratios 3. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/&sa=D&source=editors&ust=1725441104802031&usg=AOvVaw3v5f6sjG8WDrN52MQJkFhG 4. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 5. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection	
6	Skin Analytics	3.4	We are concerned that the implicit focus on teledermatology and lack thereof on face-to-face care overlooks the workforce limitations that underpin the unmet need: - 24% of Dermatology consultant positions remained vacant, with only 508 whole-time-equivalent (WTE) consultants in posts and 159 WTE vacancies [8]; - Despite recommendations to increase training posts, the number fell to 32 across England in 2023 [9] - In 2023 there were >680,000 urgent suspected skin cancer referrals, which is increasing 11% year-on-year [10] 1. https://doi.org/10.1002/14651858.CD011902.pub2 2. https://doi.org/10.1002/14651858.CD013193 3. https://doi.org/10.1038/s41746-024-01103-x 4. https://doi.org/10.1038/s43856-024-00598-5 5. https://doi.org/10.3389/fmed.2024.1302363 6. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf 7. https://skin-analytics.com/ai-pathways/derm-performance/ 8. https://gettingitrightfirsttime.co.uk/wp-content/uploads/2021/09/DermatologyReport-Sept21o.pdf 9. https://medical.hee.nhs.uk/medical-training-recruitment/medical-specialty-training/competition-ratios	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.

			10. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/&sa=D&source=editors&ust=1725441104802031&usg=AOvVaw3v5f6sjG8WDrN52MQJkFhG	
7	Skin Analytics	3.9	This statement suggests that the committee has misunderstood the role and potential benefits of an AI triage. Whilst a teledermatology service would reduce the number of face-to-face dermatologist appointments, that teledermatology review still requires dermatologist time. It is well recognised that there are not enough dermatologists available and that demand outstrips that capacity. There are no additional dermatologists ready-in-waiting and specialist training requires at least seven years; therefore the autonomous use of DERM offers the only viable option to increase capacity. Moreover, this is a misunderstanding of the pathway. DERM assessment occurs before a teledermatology review. DERM assessed patients are discharged or routed to Trust teledermatology. The Trust teledermatologist can discharge patients during their subsequent review also. Nevertheless, the DERM-005 study compared DERM with the most well-established teledermatology service in the country (since 2017), Chelsea & Westminster Hospital, and found comparable sensitivity and greater specificity of DERM [1]. They have since implemented DERM and have seen substantial capacity gains including a 67% improvement to their conversion rate. 1. https://doi.org/10.3389/fmed.2024.1302363	Thank you for your comment, which the committee has considered. The committee noted that neither DERM 003 nor DERM 005 did a test of equivalence, so it is not certain that DERM has equivalent sensitivity to teledermatology and face-to-face dermatologist review for detection of cancer lesions. But, the confidence intervals overlapped in both studies, and the committee agreed that there is no evidence to suggest that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identification of cancer lesions. The EAG reported that if the sensitivity of automated DERM is 95% then use of a healthcare professional review could increase this to around 98%, based on data from Edge Health's Evaluating AI Implementation in the NHS report. The committee acknowledged that using DERM with a healthcare professional review could reduce the risk of missing skin cancers, but it was uncertain of the impact of this approach on dermatologist capacity (see section 3.11). The committee concluded that further evidence should be generated on the sensitivity of automated DERM to detect cancer lesions used within a well-established teledermatology service compared with the

				sensitivity of a well-established teledermatology service alone. See section 3.5 of the guidance. Higher specificities would result in higher discharge rates of non-cancer skin lesions. Thomas et al. (2023) reported specificities for detection of cancer lesions ranging from 70.1 to 73.4% with using automated DERM at 2 NHS trusts. The second read reviewer overturned 40% to 50% of the cases that DERM had marked as eligible for discharge. Marsden et al. (2024) reported specificities for detection of cancer lesions of 73.3% (95% CI: 69.9 to 76.4) for automated DERM in a real-world setting and 71.9% (95% CI: 68.4 to 75.1) for teledermatology. Marsden et al. (2023) reported the specificity for cancer detection with automated DERM to be 45.0% (95% CI: 39.5 to 50.6), which was considerably lower than the 77.4% (95% CI: 72.4 to 81.8) for face-to-face dermatologist review. These results suggest that DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the

				capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
8	NHS England	1.2	NHSE is keen to oversee the controlled adoption of this technology as this is part of a wider NHSE response to system risks. For example, this includes (1) the many patients with melanomas and SCC who are incorrectly placed by primary care onto routine waiting lists and (2) Those coming to harm whilst waiting.	Thank you for your comment, which the committee has considered. The committee heard the company's and NHS England's suggestions on managing risks and agreed that it would be important to mitigate potential risk while DERM is used within clinical practice. This could be done through using a healthcare professional review, safety net protocols to prevent missed or delayed cancer diagnoses, regular monitoring of DERM's accuracy, and a national governance framework to ensure local oversight of use of DERM. See section 3.16 of the guidance.
9	NHS England	1.3	There is variation in the uptake of TD and AI technologies across providers making comparative data collection very difficult. Given the strains on the current service and predicted growth of demand, any optimisation of clinical capacity is warranted. AI's ability to safely triage benign lesions away from the USSC without the need for dermatologist review, will benefit all patients on waiting lists due to the release of clinical capacity which can be redeployed for accelerated reduction of these lists.	Thank you for your comment, which the committee has considered. The committee discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.
10	NHS England	3	NHSE request that committee remain especially cognisant of the current scale of unmet need in the UK. Incoming referrals onto skin pathways represent quantities of work which are equal to or greater than the existing	Thank you for your comment, which the committee has considered.

			clinical workforce's routine capacity to review and treat patients. Large proportions of USSC appointments, be they F2F or TD, which result in discharge with no further treatment are a suboptimal use of highly valuable members of the clinical workforce. These appointments should be safely minimised in order for patients and services to maximally benefit from these individuals' expertise. This technology presents an opportunity to streamline the clinical capacity required for initial assessment of patients on the USSC, thereby freeing up scarce specialist clinical resource for more impactful work with patients who have malignancies and would otherwise have had to wait longer for their cancer to be treated, and/or for patients with other inflammatory skin conditions and for whom current treatment capacity is extremely limited. 2024 NDRS data suggest that as many as 48% of malignancies may be identified via the routine pathway; indicating significant patient harm from routine assessment delay.	The committee acknowledges there is an unmet need and these points have been included in section 3.1 of the guidance.
11	NHS England	3.5	The benefit of autonomous AI technologies is the release of workforce, and improved and timely treatment of malignant lesions.	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and section 3.1 of the guidance notes that AI technologies used within a teledermatology service could potentially increase staff capacity in dermatology services to help address the unmet need.
12	NHS England	3.9	This statement fails to account for the fact that the TD clinic itself is a use of the same clinical workforce which would be conducting the F2F assessments, and therefore the overall productivity gain remains limited. Only the automated AI can genuinely reduce the clinical workforce burden posed by initial assessments of suspicious lesions	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and section 3.1 of the guidance notes that AI technologies used within a teledermatology service could potentially

				increase staff capacity in dermatology services to help address the unmet need.
13	Consultee 4	3.1	It is clear that there are not enough dermatologists to ensure face to face appointments within a reasonable time scale. It is also clear that delayed appointments can lead to delayed diagnosis and potential death through melanoma. It is therefore urgent to present alternative solutions other than F2F and DERM and AI must be a part of this solution frreing up viatal resource to help address the unmet need.	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and section 3.1 of the guidance notes that AI technologies used within a teledermatology service could potentially increase staff capacity in dermatology services to help address the unmet need.
14	Healthcare professional 1		When considering the unmet need it must also be highlighted that at least a third of melanomas and SCCs are found on routine pathways. Prioritisation of cancer pathways, backlogs for new routine referrals and follow up for high risk patients with cancer all mean that cancers are being picked up later and with worse outcomes. In addition, lack of capacity for minor operations means that treatment is often delayed. Our Trust was the first to establish a teledermatology service for urgent suspected skin cancer in the UK and is now used as a national exemplar for outpatient recovery. The teledermatology service has been running for 7 years during which time it has been honed and streamlined to increase efficiency. Despite this we were unable to keep up with the increase in urgent cancer referrals (175% increase in 2021) and a growing routine waiting list (>4000 patients) for which we have had to resort to costly private insourcing which is unsustainable. There are no other robust solutions to this issue and we feel confident that the AI technology can provide us with a safe way to increase capacity allowing us to make the very best use of dermatologist time directing it to where it is most needed.	Thank you for your comment, which the committee has considered. The unmet need has been noted in section 3.1 of the guidance.

Theme: Research use only recommendation

Comment number	Name and organisation	Section number	Comment	NICE response
15	Birmingham AI and Digital Health Research Team, University of Birmingham	1	Thank you to the NICE EVA committee for these recommendations in addition to the extensive report. Please accept my comments on behalf of the Birmingham AI and Digital Health Research Team. Overall we agree that further evidence is required, however we believe that (for many use cases) the evidence generation process should extend into the post-deployment phase. This is particularly relevant for generation of evidence in darker skin tones.	Thank you for your comment, which the committee has considered. The final guidance recommends that DERM can be used within teledermatology services in the NHS during the evidence generation period as an option to assess and triage skin lesions in adults referred to the urgent suspected skin cancer pathway. The evidence generation plan gives further information on the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.
16	Consultee 2	1	The recommendations are incorrect. Adequate research has already been done answering the outstanding questions posed for DERM and it is highly debatable if further research would reduce uncertainty. It would also be costly, potentially unethical and delay the introduction of a technology which has been demonstrated to be effective	Thank you for your comment, which the committee has considered. See response to comment 15.
17	Consultee 2	1.2	The four areas of research are unnecessary. My further comments explain why for each area. In some cases the research is unnecessary because there is already reasonable certainty about the question posed. In other cases the research proposed is not feasible	Thank you for your comment, which the committee has considered. See response to comment 15.
18	Skin Analytics	1	There are a number of material misunderstandings on how DERM works within the pathways it has been deployed within, which appear to underpin some of the evidence gaps proposed.	Thank you for your comment, which the committee has considered.

			The requirement to collect additional data for the performance of DERM in black and brown skin raises an important question to help reduce health inequalities. However, the current standard of care has significant health inequalities and a risk benefit evaluation should identify whether it is ethical to deny access to DERM for an indefinite period to collect the data, especially when this continues to be monitored in a thorough post market surveillance programme. Such a requirement would deny access to a highly sensitive technology which could significantly reduce the requirement for dermatologist appointments in the urgent cancer pathway (both points made by the analysis) which could be used to clear routine waitlists. Please see our comments, particularly in section 1.3, for all of the specific data points pertaining to our points raised above.	See response to comment 15.
19	Skin Analytics	1.3	This paragraph does not reflect the 3 health economic assessments (1 commissioned and co-created by Skin Analytics) of DERM pathways which were submitted and the further report since published by Edge Health commissioned by NHSE that consistently show circa £2 benefit for £1 cost [1-4]. 1. https://openresearch.surrey.ac.uk/esploro/outputs/report/99873466002346 2. https://www.edgehealth.co.uk/news-insights/evaluation-nhs-ai-skin-cancer/ 3. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf 4. https://skin-analytics.com/wp-content/uploads/2023/12/Skin-Analytics-Health-Economics-impact.pdf	Thank you for your comment, which the committee has considered. See response to comment 15.
20	NHS England		NHSE Overarching comment: There is a complicated landscape in the delivery of dermatology services where need outstrips capacity and where alternative models of interventions would be helpful. There is also much divergence of thinking within the dermatology community but the following is a unified view from a large range of clinicians consulted by NHSE. NHSE believe that the recommendation should be changed to: "Shows sufficient	Thank you for your comment, which the committee has considered. See response to comment 15.

			promise to conditionally recommend for use with data collection; conditionally recommend for use while further evidence is generated" NHSE have reached a consensus on this based on the below points of clinical evidence and assessed operational capability: • NICE have used a sensitivity analysis to capture the diagnosis of a malignancy but there is a larger opportunity to exclude benign pathology. This would free up clinical capacity to both treat those patients with malignant disease as well as improve access for those on a non-2-week pathway. This would require an analysis of a negative predictive value which was not addressed in this EVA. • The recommendation needs to be adjusted in order to enable a non-research model to be deployed safely by implementing a nationally steered prospective service evaluation with integrated prospective data collection across all NHS sites • This service evaluation would be paired with an alternative human-led clinical review whilst prospective data is gathered nationally on the most recent software. This would continue until it is clear to the main stakeholders that it is safe enough to roll out autonomously with consent and data gathering but without a clinical study wrap around. • In considering appropriate deployment, with clinical governance including informed consent, clinical guidance, human clinical review, and mandatory data collection, the NHS must consider and balance the need to minimise harms to patients on waiting lists with unmet needs. • It is to be noted that the MHRA require the company to provide regular post-marketing surveillance data to the sites including diagnostic performance and this will also be used to test and re-test the safety of the device in terms of assessment of performance of the NPV of the tool. In addition NHSE will seek to formalise the need for the company to report data to NHSE under contract if the product were to be procured nationally.	
21	NHS England		Regarding the question: "Are the recommendations sound, and a suitable basis for guidance to the NHS?"	Thank you for your comment, which the committee has considered.

			The recommendation should be changed to 'shows sufficient promise to conditionally recommend for use with data collection'. The ongoing service evaluation and robust PMS being proposed by NHSE will ensure the 'highlighting of any risks which could be mitigated in evidence generation' and 'identify key evidence gaps to be addressed by evidence generation' which are linked to this statement in the NICE decision making flow chart. The latter enables us to capture the prospective data for the richer skin tones which is currently limited (due to these groups representing only 3-4% of presentations and ~0.5% of skin cancer malignancies in the UK, as outlined by NICE) whilst not disadvantaging those with paler skin in whom skin cancer is much more common. It would avoid unnecessary additional harms to patients on long routine waiting lists who are currently unable to access treatment in dermatology services.	See response to comment 15.
22	NHS England	1.1	It is suggested that the following recommendation is used: - 'Shows sufficient promise to conditionally recommend for use with data collection; conditionally recommend for use while further evidence is generated' The views of the Dermatology community are divergent regarding the use of Teledermatology as well as DERM but the unified position is that the uncontrolled rollout of the autonomous AI tool is not appropriate. There is absolute agreement that guardrails should be placed as this technology becomes embedded in clinical practice. This includes informed consent, human clinical review (in a service evaulation rather than research study setting) as the technology is deployed, and mandatory data collection. Current deployment of this technology in commercial or nationally funded programme contexts is supported by safety netting in a similar manner to discharge after a face to face dermatology assessment Local and national data collection and post market surveillance of this usage provide robust additions to the existing evidence base and fulfill NICE's recommendation for additional data without limiting the benefits being exprienced in those trusts which seek to deploy this technology autonomously after a risk assessment and review with a human clinical review.	Thank you for your comment, which the committee has considered. See response to comment 15.

23	Consultee 5	3	It seems to me a "no brainer" that AI should be used to quickly filter out non-cancerous skin legions and to fast-track potentially dangerous legions to be tested.	Thank you for your comment, which the committee has considered. See response to comment 15.
24	Consultee 5	3.1	Exactly, so why not go ahead now? If money is the answer, surely this would be cheaper than dealing with higher stage melanoma down the line, following the current NHS delays? Not to mention the distress caused to patients.	Thank you for your comment, which the committee has considered. See response to comment 15.
25	Consultee 5	3.1	Surely now is the time to act, rather than seek "further research". How much information is sufficient information? People are suffering right now...	Thank you for your comment, which the committee has considered. See response to comment 15.
26	Healthcare professional 3	3.3	In summary whilst understanding that data collection is most definitely needed I believe, based on our experience, that this could be completed pro-actively in secondary care as the new approach to care delivery is introduced. In all units making the move I would recommend there are clear rules for continuing systematic and standardised close monitoring. Although I cannot be certain on the absolute differences in risk between conventional and AI-supported approaches our pilot suggest these are likely to be modest. Intuitively overall I feel that net benefit is likely as the current delays are clinically unacceptable. Furthermore with continuing technological advance AI-empowered analytic approaches are bound to move to standard of care as demands on services continue to rise. So, it would be good if dermatology services are allowed to introduce AI options now if local clinical champions can be identified.	Thank you for your comment, which the committee has considered. See response to comment 15.
27	British Association of Dermatologists		AI has huge potential to offer to help remodel how Dermatology Services could be delivered in future, integrated with the HCPs to shape and provide the essential human element of care. The current EVA risks proposing a model of care without thorough evaluation of impact across the patient pathway	Thank you for your comment, which the committee has considered. See response to comment 15.

Theme: Mitigation of risk

Comment number	Name and organisation	Section number	Comment	NICE response
28	Skin Analytics	1.3	In summary, the process outlined for the EVA states the requirement for conditional approval is dependent on the balance of the risk and benefit from use of the intervention. Without the data outlined above, it is not clear that this process has been followed.	Thank you for your comment, which the committee has considered. The final guidance recommends that DERM can be used within teledermatology services in the NHS during the evidence generation period as an option to assess and triage skin lesions in adults referred to the urgent suspected skin cancer pathway.
29	NHS England	1.3	The second read requires dermatology capacity and may not be more efficient than a trust teledermatology clinic but will be compared to F2F assessments. A human clinical review for a limited time period with sufficient data collection when DERM is initially embedded within a service is an appropriate mitigation to a percieved risk when a service is first started up. Trusts may gain productivity from DERM use which includes a human clinical review during this initial period. The model for temporary human clinical review may need to expand from the current SkinAnalytics in house second read to a much wider base to support this need.	Thank you for your comment, which the committee has considered. The committee acknowledged that using DERM with a healthcare professional review could reduce the risk of missing skin cancers, but it was uncertain of the impact of this approach on dermatologist capacity (see section 3.11). The committee concluded that further evidence should be generated on the sensitivity to detect cancer lesions of automated DERM used within a well-established teledermatology service compared with the sensitivity of a well-established teledermatology service alone, as well as the impact this has on dermatologist capacity.
30	NHS England	1.3	The national recommendations are important here as to how to safety net and recall as needed to reduce the risk of missed or delayed cancer diagnosis. NHSE recommends appropriate, regular monitoring of the performance of DERM technology to mitigate against the potential possibility of a dip in	Thank you for your comment, which the committee has considered. The committee acknowledged these strategies to manage risk and information has been included in section 3.16 of the guidance.

			accuracy. The datasets are discontiguous in providing updates into the software providing a real time opportunity to understand if the tool becomes inaccurate.	
31	NHS England	2.1	It should be noted, as per previous comment, that the device is updated discontiguously and updated versions are released to services periodically. As per previous comments, this version control combined with NHSE monitoring provide an opportunity to compare its performance across updates to ensure maintained safety and effectiveness. This will further enhance confidence.	Thank you for your comment, which the committee has considered. The committee acknowledged these strategies to manage risk and information has been included in section 3.16 of the guidance.
32	NHS England	3.2	TO ensure that any services which are due to deploy are content with this device in their service, NHSE proposes that there will be local governance within a national framework as to the adoption of a new technology within a provider with sign off of clinical risk and assurance. Deployment will be a trust decision which comes with the requirement to act within nationally overseen assurance processes.	Thank you for your comment, which the committee has considered. The committee acknowledged these strategies to manage risk and information has been included in section 3.16 of the guidance.
33	NHS England	3.13	As per previous comments, NHSE agree and there are mitigations and recommendations in place which recognise this.	Thank you for your comment, which the committee has considered. The committee acknowledged these strategies to manage risk and information has been included in section 3.16 of the guidance.

Theme: Healthcare professional views

Comment number	Name and organisation	Section number	Comment	NICE response
34	Skin Analytics	3.3	This section does not take into account all of the relevant evidence. For example, it does not acknowledge the 18 provider staff interviews conducted in the AI in Health & Care Award evaluation which was submitted and included findings that DERM "allowed for the discharge of patients with benign conditions at triage, increasing capacity for those with more complex needs" and "significantly improved dermatology waiting times and capacity" as well as a desire for higher quality images for the Trust teledermatology review [1]. In addition, further evidence has recently been published with the NHSE-commissioned Edge report including provider interviews. These included findings of: "reduction in patient reviews and F2F appointments", being able to "meet and surpass cancer and performance targets, despite an increase in referral numbers", "a reduction in their biopsy rates", feedback that "patients expressed that they could be seen more promptly, reducing the need for multiple hospital visits", as well as challenges such as Wi-Fi connectivity, system integration and more complex cases for face-to-face consultations [2] Without appropriate acknowledgement of this evidence, it is possible the committee have not been fully informed on the potential and realised benefits and acceptability to healthcare professionals. Should this be considered it would support conditional recommendation, in line with the EVA decision making flow diagram seen in committee papers 2. 1. https://openresearch.surrey.ac.uk/esploro/outputs/report/99873466002346#file-0 2. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf	Thank you for your comment, which the committee has considered. The committee acknowledged these observations and they have been included in section 3.3 of the guidance.
35	Unity Insights and University of Surrey	3.2	Some clinicians felt strongly that the second read was an important measure in ensuring the safety of the pathway and were nervous at the prospect of this being removed, which is unsurprising given that this technology is still relatively new. However they did not rule out removal of the second read as the technology matured with the accrual of more data.	Thank you for your comment, which the committee has considered.

				The committee acknowledged these observations and they have been included in section 3.3 of the guidance.
36	Unity Insights and University of Surrey	3.2	· We interviewed 18 members of staff to find out more about their opinions on the use of AI technologies. · We found that establishing teledermatology services using DERM has had a significant and positive impact on dermatology waiting times and capacity. Staff felt that it has been 'transformational' to dermatology services. It is difficult to establish the extent to which this is a result of DERM specifically or teledermatology more generally. However, they identified that there is inherent risk in reviewing patients remotely as opposed to F2F associated with not being able to conduct a thorough and holistic clinical examination. · One of the main benefits of DERM was perceived by staff to be its ability to discharge patients with benign disease at triage, thus increasing capacity in the dermatology service. Staff from all sites proposed that, in the long-term, DERM would be best used as a triage tool in primary care. In this model, benign lesions could be referred back to the GP, and suspicious lesions would proceed to a dermatology review in secondary care involving DSLR images and/or a face-to-face appointment. · We identified a number of important factors involved in setting up the DERM service successfully which included: Quality of referral, reliable internet connectivity, support and training for staff, high-quality images and patient information.	Thank you for your comment, which the committee has considered. The committee acknowledged these observations and they have been included in section 3.3 of the guidance.
37	Healthcare professional 3	3.3	At our Trust we have been using DERM in the post-referral setting for USSC referrals since October 2021 as we cope under considerable strain with the huge increase in referrals since COVID. Our unit is based in rural East Anglia where predicted lifespans are long, people have often been employed outside and cloud cover can be low. The discharge rate in our initial pilot experience with DERM has been approximately 20% without requiring Trust input. 66% of USSC cases have avoided a 2ww F2F appointment and instead have been booked with other specialties and timelines. To date we have entered almost 9,000 cases with data being collected continuously. We continue to closely monitor the service highlighting checks for safety especially seeking out any	Thank you for your comment, which the committee has considered. The committee acknowledged these observations and they have been included in section 3.3 of the guidance.

			potential missed diagnosis of skin cancers. I believe Skin Analytics have been fully transparent in sharing ongoing data and we have had quarterly meetings to discuss our practical experience, the emergent data with a hard focus on any false negatives. These false negative i.e., missed cancers are then subject to deep dive review and root cause analysis. To date there has been a single missed lentigo maligna, two missed SCC's and one missed BCC. At our Trust we have had the support from IT and Legal & Governance Departments using DERM. Patient feedback has been positive. As from early September 2024 we have decided that our Trust Dermatologists will no longer perform the second read for benign lesions and are now using the AI autonomous pathway for benign lesions so that they have more time to concentrate on possible skin cancer. We continue with intensive post-market surveillance assessing sensitivity and specificity of the AI service. In summary whilst understanding that data collection is most definitely needed I believe, based on our experience, that this could be completed pro-actively in secondary care as the new approach to care delivery is introduced. In all units making the move I would recommend there are clear rules for continuing systematic and standardised close monitoring. Although I cannot be certain on the absolute differences in risk between conventional and AI-supported approaches our pilot suggest these are likely to be modest. Intuitively overall I feel that net benefit is likely as the current delays are clinically unacceptable. Furthermore with continuing technological advance AI-empowered analytic approaches are bound to move to standard of care as demands on services continue to rise. So, it would be good if dermatology services are allowed to introduce AI options now if local clinical champions can be identified.	
38	Healthcare professional 1		Our Trust was the first to implement a teledermatology service for urgent suspected skin cancer which we did in 2017. In the interim we have optimised the service which performs well and has been awarded the BMJ dermatology team of the year and used as an exemplar by NHSE. Despite this urgent skin cancers have continued to increase with a	Thank you for your comment, which the committee has considered.

			175% increase in referrals in 2021 and a growing backlog of new and follow up routine appointments. As a result we were unable to keep up with demand and the natural progression was to implement DERM which we deployed >2 years ago. Since that time we have reviewed over 12000 patients on this pathway. 94% of patients avoided an urgent F2F appointment, we saw a 10% reduction in biopsies, a 13% reduction in routine follow up appointments and rates of urgent appointments were unchanged. 37% of cases were suitable for immediate discharge without clinician review. Together this has provided us with much needed capacity as our urgent cancer referrals continue to increase circa 20% per year. In addition following implementation of DERM the conversion rate for skin cancer increased by 67%. A survey of dermatologists at our trust revealed that 100% had a positive experience with the DERM software and that all would be disappointed if they could no longer use the technololgy. Since deployment in April 2022 DERM has missed 1 melanoma in situ, 2x pT1a melanomas, 2x SCCs and 1 rare cancer. In comparison, anecdotally in the last 8 months on non-AI pathways we have noted missed or delayed diagnoses for 6 invasive melanomas, 5 MIS, 1 angiosarcoma, 1 sebaceous carcinoma, 4 SCCs and 1 metastatic merkel cell carcinoma. The reality is that data on clinician and pathway sensitivity is woefully inadequate and it is likely that we overestimate performance in these settings. 3.4 See above comment in 3.3.	The committee acknowledged these observations and they have been included in section 3.3 of the guidance.
39	Healthcare professional 2		HCA delivering the photography depends on the calibre of the HCA – but SA have offered excellent training, and ongoing support. They are responsive. My own feeling is that we are 6 months into our trial and I see it overcalling BCC that are ulcerated as MM. The HCA may not be clinically skilled to recognise an ulcerated lesion as unsuitable for analysis and so this data is entered into the training bank of the algorithm- I feel that when there is a mismatch between Clinician reviewing patient after a DERM AI outcome that stated high risk, it is difficult to advise the patient that no treatment is required and so in our service I feel it is likely biopsy rates have increased not decreased as it overcalls – as set with low threshold to go with diagnostic code of greatest severity if it is ambiguious.	Thank you for your comment, which the committee has considered. The committee acknowledged these observations and they have been included in section 3.3 of the guidance.

Theme: Patient views

Comment number	Name and organisation	Section number	Comment	NICE responses
40	Birmingham AI and Digital Health Research Team, University of Birmingham	1.3	As discussed above, we agree that further evidence is required due to small sample sizes. We delivered a PPIE workshop discussing the use of AI for skin cancer and raised the issue of differential performance across skin tones. Most of the patients were keen for the technology to still be deployed, however there should be a closely monitored pathway for patients with darker skin tones at least during the implementation process and for a duration afterwards.	Thank you for your comment, which the committee has considered. The committee emphasised that the amount of data remains small. So, more data is needed on the accuracy of automated DERM in people with black or brown skin to be sure it does not incorrectly detect (false positive) or miss (false negative) skin cancer. It also noted that it is important to use AI technologies with a healthcare professional review in this group until more data is available. See section 3.9 of the guidance.
41	Birmingham AI and Digital Health Research Team, University of Birmingham	3.2	Although this concern is understandable, based on the evidence available it seems that the sensitivity of DERM is higher than dermatologists alone (although non-significant based on 95%CI). With regards to the safety net advice, this is usually given in a routine GP appointment anyways. I believe these concerns will be better responded to with patient education and appropriate resources.	Thank you for your comment, which the committee has considered. The committee acknowledged the emphasis on patient education. Section 3.2 of the guidance has been updated to state "Stakeholders and the clinical experts highlighted that education for people using the service and clear communication are important in building peoples' trust in AI technologies. Education should provide people using the service with enough information on the waiting times and performance of AI compared with face-to-face and teledermatology assessments. This will allow people to make an informed

				decision on whether to accept or decline the use of AI technology".
42	Consultee 2	3.2	The patient concerns about the safety of the technologies (missing cancer diagnoses) are very reasonable. However, whether they are justified depends on how well the evidence has been explained to patients. As will be explained later the best evidence available suggests that the DERM misses diagnoses at a similar level to current practice, whether that is teledermatology or face-to-face consultation. Patients may be believe that current practice never misses cancer diagnoses, but the evidence suggests this is not the case.	Thank you for your comment, which the committee has considered. See response to comment 41.
43	Consultee 1		Patient Consideration 3.2:- Patients indeed may be concerned at not having a face to face consultation, but the proven level of 99.85% accuracy of excluding non malignant lesions via AI as established by Skin Analytics is truly astonishing. Especially as the result is available within 10 seconds of the non invasive dermal scan. For non complicated, as with a previous history or older/multiple lesions, then the current Dermatology Consultant face to face is the way to go, to alleviate any patient concerns. That does therefore open the AI scan pathway to the vast majority of first referrals, so dramatically reducing the wait list for urgent consultations requiring face to face appointments.	Thank you for your comment, which the committee has considered. See response to comment 41.
44	Skin Analytics	3.2	We would ask for this section to better reflect the volume of data and patient opinions that were sought and fed into the evidence submitted relating to DERM pathways. 1,180 survey responses and over 70,000 consent responses from patients fed into the 1 research study [1], 2 independent service evaluations [2,3] and 1 company report [4] submitted. The broad findings of all of these reports have been consistently that patients are amenable to the use of AI in their skin cancer care, particularly if the alternative is waiting weeks for a dermatologist appointment. The latter is key given the capacity constraints dermatology services face. We want to ensure the scale of these patient perspectives data are appropriately considered alongside those of the committee members. 1. https://doi.org/10.3389/fmed.2023.1259595 2. https://openresearch.surrey.ac.uk/esploro/outputs/report/99873466002346	Thank you for your comment, which the committee has considered. Findings from these reports have now been included in section 3.2 of the guidance.

			3. https://www.edgehealth.co.uk/news-insights/evaluation-nhs-ai-skin-cancer/ 4. Patient Sentiment Report submitted by Skin Analytics in Oct and Nov 2023	
45	Skin Analytics	3.2	It is unclear as to why this would differ from teledermatology. It is crucial that patients understand all of the information presented to them throughout any clinical pathway. To help us achieve this, we have worked with patients and relatives, including those from the support group MelaNoMore, to carefully create all of our patient facing materials. We have received positive feedback from partners that the content in our templated letters surpasses what they had previously been using in their pathways. In addition to our efforts to engage regularly with patients, we know that a number of our NHS partners co-create their letters and information leaflets with local Patient & Public Involvement groups prior to launching DERM pathways.	Thank you for your comment, which the committee has considered. See response to comment 41.
46	NHS England	3.2	While patients' feelings should hold weight, it is the role and challenge for organisations such as NICE and NHSE to progress with the development of treatments and interventions which are supported by evidence. Our role as healthcare professionals is to communicate with patients in an evidence led way rather than allow sometimes incorrect perceptions from laypersons to wrongly influence which technologies are offered. The acceptance rate of DERM in the real world is very high. It is a statutory part of the service that all patients are given sufficient information to give informed consent. Patients are also given the option of rejecting autonomous assessment via AI in favour of joining a F2F or TD waiting list. At a trust where this service is currently being used autonomously, 99.95% of patients to date have consented to autonomous AI assessment on this basis.	Thank you for your comment, which the committee has considered. See response to comment 41.
47	NHS England	3.2	This is similar to many other forms of healthcare services. Dermatology services face two dilemmas: identifying and treating those with a malignant lesions in a timely fashion; and supporting patients with benign disease. By identifying those with benign disease, there is a further opportunity to provide face to face or online support via group consultations etc using a wide multiprofessional dermatology team.	Thank you for your comment, which the committee has considered. Section 3.2 of the guidance has been updated to note that the people interviewed [for the Unity Insights report] expressed high levels of satisfaction with the service, particularly because of how

			Many people prefer the idea of a teledermatology interaction rather than F2F with a reassurance and advice for the benign condition.	quickly they were assessed. For some people, the reassurance of a healthcare professional review was important to their acceptance of AI.
48	Unity Insights and University of Surrey	3.2	Context regarding the independent evaluation qualitative study conducted: Participants: 478 participants across 4 sites completed our survey. We interviewed 39 people: · 18 members of staff (5 clinical photographers, 7 dermatologists, 5 management/administrative staff, 1 HCA) · 21 patients Observations/comments regarding this section: · Patients reported high levels of satisfaction with the service, particularly in terms of the speed with which they are seen and the interaction with staff at the teledermatology appointment. When things didn't go well this was generally associated with a breakdown in communication (e.g. not informed of appointments, confusion over letters etc.) · Preference for F2F appointments may be linked to expectation setting and communication. Sites noticed an improvement in patient satisfaction when information was refined so that patients knew what to expect beforehand. · Patients were largely optimistic about the use of AI in the form of DERM. We received a number of extremely positive responses, with some patients expressing excitement at the potential use of this technology. They recognised that the NHS was under pressure, and that this approach seemed to offer a desirable solution. · It is important to note that patients were generally informed that their images would be reviewed by a dermatologist as well as DERM. For some patients, the reassurance provided by the second read was important to their acceptance of the use of AI in their care.	Thank you for your comment, which the committee has considered. See response to comment 47.

49	Consultee 5	3.2	Clearly this is true in an ideal world, but as noted above, there is currently insufficient capacity in the NHS. Surely DERM should filter where possible, freeing up the dermatologists to deal with the most serious cases?	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and section 3.1 of the guidance notes that AI technologies used within a teledermatology service could potentially increase staff capacity in dermatology services to help address the unmet need.
50	Consultee 4	3.2	I not that the two lay members of this committee had a preference for face to face dermatology appontments. What is not discussed here is the duration to obtain a F2F appointment and the wait time for a subsequent biopsy diagnosis .Locally for me a F2F can take up to 6 weeks and if needed surgery another 4 week wait and diagnosis of a biopsy 10 weeks. Anything to reduce this wait time is welcomed. Given a choice of waiting 6 weeks to get a F2F or using say DERM withih a week I would always go for the latter. There are a lot of stats in this report which seem to be delaying implementation as the report says more needed. There is no indication as to how much more is needed! From the stats I would be comfortable using DERM if it reduced the wait time. It would appear to me that this report is one of prevarication and should be much more proactive in defining what and quantifying what needs to be done to enable implementation of these technologies. If one has an advantage over another then split the report or state the minimum thresholds the technology has to meet. The object has to be to implement the technology in the quickest time scale in order to alleviate the pressure on the dermatologists and to reduce patient wait times and recommendations should be proactive in making this happen. Angst in patients is real especially if you have a suspect mole. I alway remember an elderly couple in a waiting room. He was crying and being consoled by his wife and sobbing to her that they cant see me for another 6 weeks and then I have to wait another 10 weeks for a diagnosis. I don't want to die. Yes angst is very real and should be taken seriously. If DERM + AI relieves angst	Thank you for your comment, which the committee has considered. See response to comment 47.

			then this report needs to help technologies to reach the required goals for implementation and not be an administrative report that ticks a box.	
51	Healthcare professional 1		Where is the evidence to support the fact that many people want a face to face appointment? This would of course also be true of teledermatology irrespective of AI and sadly is not achieveable in the current climate. The correct question to ask the lay members would be whether they would rather have their skin lesion assessed in 1-2 weeks or wait several months for a face to face review. Our patient satisfaction surveys have revealed that most respondants had no preference whether they were seen face to face or assessed by a computer. In addition it is important to highlight to the lay members that misdiagnosis and delayed diagnosis still occur on conventional face to face pathways, which we know is highly variable depending on where you are seen. The Edge report demonstrates that AI performance is at least as good (if not better) than face to face and teledermatology review and therefore provides much better standardised care. In addition offering safety netting advice is best practice for Face to face as well as teledermatology and AI pathways and so this issue is not unique to AI technology (although with the latter it can be standardised so that all people receive the same level of information).	Thank you for your comment, which the committee has considered. See response to comment 47.
52	Healthcare professional 2		Despite good PIL documents – many patients who attend are not fully informed about the care pathway – so need extra time to go through CONSENt process/ understand DERM AI	Thank you for your comment, which the committee has considered. See response to comment 41.

Theme: Comparative accuracy of DERM

Comment number	Name and organisation	Section number	Comment	NICE response
53	Birmingham AI and Digital Health Research Team, University of Birmingham	1	Finally, there is limited discussion about dermatologist performance. The full report states that DERM alone performs better (sensitivity) than dermatologists alone. In the post-referral pathway, specificity is less relevant. This means that DERM is non-inferior for triaging of cancers and has benchmarked against dermatologist sensitivity. A lower specificity in the post-referral pathway means that there will still be a reduction in f2f referrals, although this difference will need to be evaluated for cost-effectiveness.	Thank you for your comment, which the committee has considered. Committee noted that Thomas et al. (2023) reported data on automated DERM used at 2 NHS trusts, which found sensitivity for detecting cancer lesions ranging between 96 to 100%. DERM 005 (Marsden et. al 2024) reported a sensitivity of 94.0% (95% confidence intervals [CI]: 84.7 to 98.1) for automated DERM in a real-world setting compared to 97.0% (95% CI: 88.7 to 99.5) for teledermatology, while DERM 003 (Marsden et al. 2023) reported sensitivities for detecting cancer lesions of 96.0% (95% CI: 92.6 to 98.0) for automated DERM compared to 93.8% (95% CI: 90.0 to 96.3) for face-to-face dermatologist review. Neither DERM 003 nor DERM 005 did a test of equivalence, so it is not certain that DERM has equivalent sensitivity to teledermatology and face-to-face dermatologist review for detection of cancer lesions. But, the confidence intervals overlapped in both studies, and the committee agreed that there is no evidence to suggest that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identifying cancer lesions. The committee concluded that further evidence should be generated on the sensitivity to detect cancer lesions of automated DERM used within a well-established teledermatology

				service compared with the sensitivity of a well-established teledermatology service alone, as well as the impact this has on dermatologist capacity. See section 3.5 of the guidance. Marsden et al. (2024) reported specificities for detection of cancer lesions of 73.3% (95% CI: 69.9 to 76.4) for automated DERM in a real-world setting and 71.9% (95% CI: 68.4 to 75.1) for teledermatology. Marsden et al. (2023) reported the specificity for cancer detection with automated DERM to be 45.0% (95% CI: 39.5 to 50.6), which was considerably lower than the 77.4% (95% CI: 72.4 to 81.8) for face-to-face dermatologist review. These results suggest that DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer pathway. Compared with face-to-face assessment, results suggest that DERM used in teledermatology services would discharge fewer non-cancer lesions from the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
54	Consultee 2	3.4	The evidence on accuracy is misrepresented. Reference to previous NICE recommendations will reveal that it has recommended many diagnostic	Thank you for your comment, which the committee has considered.

			technologies with sensitivities below that achieved by DERM. It further assumes that current practice achieves a sensitivity of 100% which is manifestly not true as testified by medico-legal claims, further supported by accuracy evidence for different versions of current practice. The comments misunderstand the main objective of DERM, which is to reduce referrals. However, this should not be at the expense of increased numbers of missed cancers and this is what research has already attempted to investigate. So far the best evidence suggests that the sensitivity of DERM (used autonomously) and current practice are equivalent. There is no evidence that the sensitivity of DERM is inferior to current practice, whether that is teledermatology or face-to-face assessment	Section 3.5 on the sensitivity of DERM has been updated. Thomas et al. (2023) reported data on automated DERM used at 2 NHS trusts, which found sensitivity for detecting cancer lesions ranging between 96 to 100%. DERM 005 (Marsden et. al 2024) reported a sensitivity of 94.0% (95% confidence intervals [CI]: 84.7 to 98.1) for automated DERM in a real-world setting compared to 97.0% (95% CI: 88.7 to 99.5) for teledermatology, while DERM 003 (Marsden et al. 2023) reported sensitivities for detecting cancer lesions of 96.0% (95% CI: 92.6 to 98.0) for automated DERM compared to 93.8% (95% CI: 90.0 to 96.3) for face-to-face dermatologist review. The committee had some concerns around the risk of bias for the reference standard in DERM 003 because only 1 dermatologist provided the clinical diagnosis used as the ground truth for non-biopsied lesions. Neither DERM 003 nor DERM 005 did a test of equivalence, so it is not certain that DERM has equivalent sensitivity to teledermatology and face-to-face dermatologist review for detection of cancer lesions. But, the confidence intervals overlapped in both studies, and the committee agreed that there is no evidence to suggest that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identification of cancer lesions.
55	Consultee 2	3.5	The Marsden et al 2024 study (on which I am a co-author - which I draw your attention to as a possible conflict of interest) is absolutely critical to the evidence base and has not been given sufficient weight. It is a comparative accuracy study of exactly the design that the recommendations suggest	Thank you for your comment, which the committee has considered. See response to comment 53.

			should be done, and arguably it already provides the evidence with reasonable certainty on how DERM performs relative to teledermatology. First this study confirms that BOTH DERM and teledermatology missed 1 of 7 cases of melanoma. This emphasises that there are skin cancer diagnoses which are challenging irrespective of whether the assessment is by DERM or an experienced dermatologist. The presence of missed lesions reflects clinical reality, not poor diagnostic performance by either AI or clinicians. 100% sensitivity is simply not achievable because of the nature of the disease. Second, the study was specifically designed and powered to test whether specificity was improved with DERM, and this hypothesis was convincingly proven. This led to the positive conclusions concerning reduction in need for face-to-face assessment with DERM. Concerning sensitivity we can be less conclusive, because the study was not powered to specifically examine this (even though the study is much larger than most accuracy studies). However, even so it provides reasonable evidence that sensitivity is equivalent between DERM and teledermatology (this is the importance of the noted overlapping confidence intervals, the significance of which should be explained in the guidance as most readers will not be familiar with the interpretation)	
56	Consultee 2	3.5	The suggestion that it would be difficult for DERM to exceed the sensitivity of teledermatology, and that this would negate the potential value of the technology misrepresents the main way in which DERM would impact on the patients and the health service. As already noted the main anticipated benefit would be to reduce referrals while maintaining the number of missed diagnoses at a reasonable level. The Marsden et al 2024 paper proves this is the case with reasonable, if not perfect certainty. I would suggest that it is unreasonable for NICE to demand complete certainty, and it is debatable whether it would be possible to repeat the study at a size of sufficient power to demonstrate equivalence of sensitivity. The study would need to be very much larger because of the low prevalence of disease. It is also debatable whether such a large study would be ethical or affordable. There would certainly be delay in introduction of a technology where there is already reasonable evidence that benefit will occur	Thank you for your comment, which the committee has considered. See response to comment 53.

57	Skin Analytics	1	We have left detailed comments in the document but thought it helpful to leave a summary. Our comments can be broadly summarised into the following key themes. We believe that there has been an error of process when evaluating the performance of DERM against the comparators stated in the Final Protocol published Nov 3 2023. Section 1.5 of that document identifies the relevant comparators as assessment by specialist dermatologists either remotely or in person. Despite this, the conclusions reached by the committee are explained by lack of comparator data with teledermatology alone. The guidance lacks reference to the existing evidence for in person dermatologist performance and section 3.5 of the draft guidance highlights that "The EAG did not systematically review the evidence on the diagnostic accuracy of teledermatology alone". For these reasons, it is not clear how that conclusion has been reached. In addition, a detailed review of the current standard of care including the risks to patients on these pathways as well as the health inequalities that exist. This is an essential part of the stated Early Value Assessment process and a significant factor in a balanced assessment of risk and benefit.	Thank you for your comment, which the committee has considered. The comparator has been clarified in section 2.7: This assessment has 2 comparators, urgent teledermatology services and urgent face-to-face secondary care dermatology appointments. Comparisons are discussed between DERM used within teledermatology services and teledermatology services alone, and DERM used within teledermatology services and face-to-face dermatology assessment alone.
58	Skin Analytics	1.3	This comment is not appropriately contextualised given the lack of analysis with respect to the comparators introduced in the Final Protocol published Nov 3, 2023. Further, the comment is likely to be misinterpreted to suggest DERM diagnostic accuracy is inferior to in-person dermatologist assessment which the data do not support. The risk of missed cancer is no greater than is seen in the meta-analyses for dermatologist assessments and an argument could be made that it is in fact lower. Given this statement has been made as justification of why the committee made the draft recommendation, we raise a concern of process. A balanced evaluation of the risk and benefit of the technology can not be made without a clear understanding of the risk of delayed or misdiagnosis	Thank you for your comment, which the committee has considered. Data from these references were presented to the committee at the third committee meeting. The committee concluded that based on the data from Marsden et al. (2024) and Marsden et al. (2023), results suggest that DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer

			and how this compares to the agreed comparators and the current standard of care. To restate data presented earlier: - demonstrated in Cochrane reviews of 92% sensitivity of in-person dermoscopy for melanoma (N=23,169) [1] and 95% sensitivity of teledermatology for skin cancer (N=717) [2]; - 84% sensitivity of teledermatology for skin cancer seen in a meta-analysis published in Nature this year [3]; and - 98.9% negative predictive value for melanoma seen in the meta-analysis conducted by Edge Health commissioned by NHS England this year [4]. - The suggestion during the previous committee meeting that dermatologist performance may be improving seems to be challenged by a new study published in Nature since this draft guidance was released that identified that dermatologist melanoma sensitivity was 73.4% (95% CI 70.1–77.0) [5]. 1. https://doi.org/10.1002/14651858.CD011902.pub2 2. https://doi.org/10.1002/14651858.CD013193 3. https://doi.org/10.1038/s41746-024-01103-x 4. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf 5. https://doi.org/10.1038/s43856-024-00598-5	pathway. Compared with face-to-face assessment, results suggest that DERM used in teledermatology services would discharge fewer non-cancer lesions from the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
59	Skin Analytics	1.3	This statement implies that to demonstrate “acceptable evidence of potential benefit” DERM would need to have superior sensitivity in detecting non-cancer to the standard of care. This appears to reflect a misunderstanding of how DERM is intended to be used. Further, we believe that the Early Value Assessment criteria to show potential benefit does not support a requirement to show superiority to standard of care for detecting non-cancer lesions when DERM is used as intended. AI technologies that are able to demonstrate non-inferior sensitivity of detecting cancer with an adequate specificity can reduce case volume and free up dermatologist capacity in a way that teledermatology alone cannot.	Thank you for your comment, which the committee has considered. The final guidance recommends that DERM can be used within teledermatology services in the NHS during the evidence generation period as an option to assess and triage skin lesions in adults referred to the urgent suspected skin cancer pathway.

			This is a benefit that has been demonstrated consistently across our pathways and acknowledged in the EAG analysis which estimated that automated use of DERM could approximately halve (47.7%) all referrals (urgent and routine combined) to a dermatologist (of lesions that can be assessed by DERM). We put forward that this is "acceptable evidence of potential benefit" which should support conditional recommendation.	
60	Skin Analytics	3.10	As per our comment in section 3.9, the DERM-005 study compared DERM with the most well-established teledermatology service in the country (since 2017) and found comparable sensitivity and greater specificity of DERM [1]. They have since implemented DERM and have seen substantial capacity gains including a 67% improvement to their conversion rate.	Thank you for your comment, which the committee has considered. See response to comment 53.
61	Skin Analytics	3.4	We would remind the committee that the standard of care in clinical pathways is for a single dermatologist to provide the clinical diagnosis and next step in a patient's care journey, including the majority of cases where no biopsy is required. DERM-003 therefore was a comparison with standard of care.	Thank you for your comment, which the committee has considered.
62	Skin Analytics	3.4	This statement is misleading when comparator data for the current standard of care is not presented. Comparator data for in person melanoma sensitivity is available and shows that at best, dermatologist in-person assessments miss 1 in 20 melanomas. When discussed by the committee was this statement considered within this context? There is substantial evidence that DERM is at least as good as dermatologists in both comparator settings described in 2.6 (teledermatology alone and face-to-face dermatologist assessment). Unfortunately this section, as written, fails to benchmark human clinician sensitivity, instead implying that dermatology pathways do not miss skin cancers. Later in section 3.5, the draft guidance highlights that "The EAG did not systematically review the evidence on the diagnostic accuracy of	Thank you for your comment, which the committee has considered. See response to comment 58.

			teledermatology alone". The guidance also lacks reference to the existing evidence for face-to-face dermatologist performance. We dispute the need for more evidence here considering the quantity of work already done and published in a number of other sources. Dermatologist performance alone: - specialist sensitivity demonstrated in Cochrane reviews is 92% (95% CI 87-95%) sensitivity of in-person dermoscopy for melanoma (N=23,169) [1] and 94.9% (95% CI 90.1-97.4%) sensitivity of teledermatology for skin cancer (N=717) [2]; - 84.2% (95% CI 76.2–89.8) sensitivity of teledermatology for skin cancer as per a meta-analysis published in Nature this year [3]; - 73.4% (95% CI 70.1–77.0) dermatologist melanoma sensitivity in a study published recently in Nature [4]. Comparison between dermatologists and DERM - The DERM-005 study compared DERM with the most well-established teledermatology service in the country and found comparable sensitivity and greater specificity of DERM [5]. - A meta-analysis conducted by Edge Health commissioned by NHS England earlier this year demonstrated "DERM achieved an NPV of 99.8% at a similar disease prevalence, demonstrating performance at least as good as that of face-to-face dermatologist evaluations." [6] DERM has consistently achieved over 95% sensitivity for cancer, with the latest post-market surveillance data comprising just under 53,000 lesions with confirmed outcomes including over 6,000 biopsy confirmed cancers [7]. The last dataset submitted to NICE on DERM performance included just under 28,000 lesions with confirmed diagnoses and demonstrated 95% (631/664) sensitivity for melanoma and 97% (3,608/3,708) sensitivity for skin cancer.	
63	Skin Analytics	3.5	This statement is not accurate and any conclusions that are derived from this statement would not be reasonably held.	Thank you for your comment, which the committee has considered.

			DERM is deployed in clinical pathways with the AI triage occurring prior to a teledermatologist review, thus offering the potential to reduce the number of dermatologist appointments compared to pathways not utilising AI triage. The effectiveness of this triage process depends on both the sensitivity and specificity of the AI tool but crucially not impacted by the sensitivity of teledermatology. In this pathway potential benefits from automated AI triage would not be impacted by the teledermatologist performance. This comment also fails to take into consideration the shortage of dermatologists to conduct these teledermatology assessments which underpin the unmet need. See previous comments on unmet need in section 3.1.	The committee noted that a well-established teledermatology service could also reduce the number of referrals to face-to-face dermatologist appointments. It discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.
64	Skin Analytics	1.3	We disagree that further evidence is required. The data submitted provides substantial evidence that DERM is at least as good as dermatologists in both comparator settings described in 2.6 (teledermatology alone and face-to-face dermatologist assessment). To summarise; dermatologist performance alone: - specialist sensitivity demonstrated in Cochrane reviews is 92% (95% CI 87-95%) sensitivity of in-person dermoscopy for melanoma (N=23,169) [1] and 94.9% (95% CI 90.1-97.4%) sensitivity of teledermatology for skin cancer (N=717) [2]; - 84.2% (95% CI 76.2–89.8) sensitivity of teledermatology for skin cancer as per a meta-analysis published in Nature this year [3]; - 73.4% (95% CI 70.1–77.0) dermatologist melanoma sensitivity in a study published recently in Nature [4]. Comparison between dermatologists and DERM: - The DERM-005 study compared DERM with the most well-established teledermatology service in the country and found comparable sensitivity and greater specificity of DERM [5]. - A meta-analysis conducted by Edge Health commissioned by NHS England earlier this year demonstrated “DERM achieved an NPV of 99.8% at	Thank you for your comment, which the committee has considered. See response to comment 53.

			a similar disease prevalence, demonstrating performance at least as good as that of face-to-face dermatologist evaluations." [6] DERM has consistently achieved over 95% sensitivity for cancer, with the latest post-market surveillance data comprising just under 53,000 lesions with confirmed outcomes including over 6,000 biopsy confirmed cancers [7]. The last dataset submitted to NICE on DERM performance included approximately 28,000 lesions with confirmed diagnoses and demonstrated 95% (631/664) sensitivity for melanoma and 97% (3,608/3,708) sensitivity for skin cancer.	
65	Skin Analytics	3.5	We note that this requirement is a corollary of an earlier point regarding the value of DERM being reliant on the sensitivity of teledermatology which is not accurate and as such the requirement for additional data must be reviewed with respect to the relative risk and benefit to patients. The DERM-005 study compared DERM with the most well-established teledermatology service in the country and found comparable sensitivity and greater specificity of DERM [1]. In addition to the data presented from DERM-005, the real world data submitted for review included over 6,500 real world cases from Chelsea & Westminster which has been running an award-winning teledermatology service. In addition, there is ample evidence baselining the sensitivity of dermatologists in a face-to-face clinical setting. There is little evidence or expert consensus that teledermatology performance would surpass the sensitivity of an in person review. DERM's performance, both in terms of sensitivity and NPV has repeatedly been shown to compare favourably. See comment on section 1.3. Without evidence or reasonable expectation that teledermatology sensitivity is significantly higher than in person assessment, it is not clear the risk to patients would support restricted use of DERM. 1. https://doi.org/10.3389/fmed.2024.1302363	Thank you for your comment, which the committee has considered. See response to comment 53.

66	NHS England	1.3	It is unclear which evidence this line refers to, as real world usage data do not indicate reduced performance amongst eligible patients when compared to other assessment options.	Thank you for your comment, which the committee has considered. See response to comment 53.
67	NHS England	1.3	Current evidence suggests that the Negative Predictive Value of DERM (absent a second read) is 99.8% at a melanoma prevalence of 2.5%. This is in comparison to an NPV of 98.9% for face to face assessment at a similar prevalence of 2.7%.	Thank you for your comment, which the committee has considered. This data was presented to the committee at the third committee meeting. The committee concluded that based on the data from Marsden et al. (2024) and Marsden et al. (2023), DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer pathway. Compared with face-to-face assessment, results suggest that DERM used in teledermatology services would discharge fewer non-cancer lesions from the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
68	NHS England	3.4	This is confusing and needs to be directly compared with the missed melanomas by a dermatologist and TD rather than stated in this way. Negative predictive value for benign vs malignant lesions is a better tool. As previously stated, the NPV for autonomous DERM is at least as high as these other methods of initial assessment.	Thank you for your comment, which the committee has considered. See response to comment 67.

69	NHS England	1.3	We believe that AI should be being compared to both face to face and teledermatology services. TD and AI are not equivalents as potential alternatives to F2F, as TD requires clinical capacity and therefore does not offer the potential for the crucial productivity gain that autonomous AI does. Moreover, F2F assessment remains the gold standard in the opinion of the dermatology community due to, for example, its ability to examine the patient more widely if deemed appropriate. Research suggests that clinical accuracy of TD clinics is marginally compromised compared to F2F clinics (NPV of 97.6% for TD compared to 98% for F2F, as outlined in the results of the systematic review conducted as part of Edge's evaluation). This implies a trade-off between efficiency and accuracy when comparing TD with F2F. Data demonstrating that AI is both more clinically efficient and has a higher NPV indicate that this trade does not exist in the use of AI.	Thank you for your comment, which the committee has considered. See response to comment 53.
70	Unity Insights and University of Surrey	3.3	Wider context around our quantitative study as part of the independent evaluation: A substantive quantitative evaluation has been undertaken by consortium comprising Unity Insights Ltd and the University of Surrey as part of the AI in Health Care Award. In total, 5,878 out of 7,447 lesions (78.8%) of all lesions referred to three secondary care sites (Chelsea and Westminster Hospitals NHS Foundation Trust, University Hospitals Bristol and Weston NHS Foundation Trust,and Ashford and St Peters NHS Foundation Trust) or considered within the community-hub at University Hospitals Birmingham NHS Foundation Trust have been assessed by DERM The full report produced is available here https://doi.org/10.15126/901081 The report confirms the performance metrics reported by Skin Analytics in their quarterly clinical performance reports, giving confidence in their monitoring and reporting in conjunction with the trusts. Accuracy of	Thank you for your comment, which the committee has considered. The committee acknowledged these observations and they have been included in section 3.3 of the guidance.

			individual lesion classification varies, but the stated benefit of DERM is through the removal of unnecessary referrals, potentially increasing the conversion rate, while maintaining a high sensitivity for cancer. There were no red flags regarding safety identified in the analysis.	
71	Unity Insights and University of Surrey	3.3	DERM is optimised for pathway sensitivity rather than sensitivity of individual lesion diagnosis; that is the ability of DERM to correctly triage referrals to the most appropriate management outcome according to risk and the locally agreed pathway. Performance: Overall, DERM yielded high pathway sensitivities for malignant melanoma across all secondary care sites (higher than 90%), and >95% for all cancers achieving its target rates, showing that DERM is effective in channelling high risk lesions to the appropriate management outcome. Published data suggests a sensitivity for all skin cancer of 94.9% from teledermatology.	Thank you for your comment, which the committee has considered. See response to comment 54.
72	Healthcare professional 1	3.5	The reality is that we don't have enough clinicians to meet the unmet need. Teledermatology whilst more efficient, is still dependent on clinician time. Where is the evidence to say that the cochrane review on teledermatology is no longer accurate? Perhaps the EAG should carry out a systematic review of teledermatology and face to face dermatology assessments if they feel the current evidence is no longer accurate. Particularly as the independent meta-analysis conducted by Edge demonstrated DERM performance was at least as good as face to face assessments.	Thank you for your comment, which the committee has considered. Data from the Edge Health report was presented to the committee. The committee acknowledges there is an unmet need, however it concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
73	British Association of Dermatologists		There is a lack of evidence of additive benefit of AI model described over routine use of a well-established teledermatology Service supported by clinical experts.	Thank you for your comment, which the committee has considered. The committee agreed that the potential benefits of DERM included:

				- DERM would not miss more cancers than a teledermatologist or a face-to-face dermatologist review of individual lesions (see sections 3.5 and 3.16) - automated use of DERM could approximately halve the number of referrals to a dermatologist within the urgent skin cancer pathway (see section 3.11).

Theme: Accuracy of teledermatology

Comment number	Name and organisation	Section number	Comment	NICE response
74	Unity Insights and University of Surrey	3.3	As noted the EAG did not examine the variation in detail, and this is a limitation of the guidance. There is clearly variation, but the guidance fails to note that suggestions that the high specificities identified in the Cochrane review are completely incompatible with routine data. If specificities were as high as 84% the number discharges after face to face appointments would be much smaller than actually observed in the routine data.	Thank you for your comment, which the committee has considered. The committee concluded that that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.
75	Unity Insights and University of Surrey	3.3	It seems highly unfair and anti-innovatory that the new technology in this appraisal is being penalised for an absence of evidence about the effectiveness of current practice, rather than absence of evidence on the technology. The accuracy of DERM is actually well characterised	Thank you for your comment, which the committee has considered.
76	Healthcare professional 1		Regarding the lack of comparator data for teledermatology, we note that in the public discussion, the Cochrane review of teledermatology performance was cited as the reference standard for that comparator and that EAG members acknowledged the limitations in this regard. We acknowledge this point and agree data on teledermatology performance is of poor quality despite widespread support for its use. The conclusion that a lack of teledermatology comparator data precludes the use of DERM given insufficient evidence of performance is not reasonable given: - This Cochrane review of teledermatology which states the data "suggest that image-based assessment may not be equivalent to a face-to-face patient:clinician interaction." [2]; - It is not reasonably held by any experts that teledermatology performance	Thank you for your comment, which the committee has considered. See response to comment 74.

			would exceed in person performance by a dermatologist; and - Comparator data exists for DERM and in person dermatologist assessments which show that DERM performs at least comparably. For there to be insufficient evidence to show promise for use within the NHS, there needs to be a reasonable expectation that teledermatology performance would be superior to in person dermatologist assessment. 1. https://doi.org/10.1002/14651858.CD011902.pub2 2. https://doi.org/10.1002/14651858.CD013193 3. https://doi.org/10.1038/s41746-024-01103-x 4. https://doi.org/10.1038/s43856-024-00598-5 5. https://doi.org/10.3389/fmed.2024.1302363 6. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf 7. https://skin-analytics.com/ai-pathways/derm-performance/ 8. https://gettingitrightfirsttime.co.uk/wp-content/uploads/2021/09/DermatologyReport-Sept21o.pdf 9. https://medical.hee.nhs.uk/medical-training-recruitment/medical-specialty-training/competition-ratios 10. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/&sa=D&source=editors&ust=1725441104802031&usg=AOvVaw3v5f6sjG8WDrN52MQJkFhG 11. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 12. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection	
77	Consultee 2	3.10	Further, while both diagnostic performance and the cost effectiveness of teledermatology both have poor quality evidence, it is not clear that the lack of this evidence presents a risk that outweighs the benefits of the capacity release that DERM enables while such data is collected. 1. https://doi.org/10.3389/fmed.2024.1302363	Thank you for your comment, which the committee has considered. The committee discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a

				well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.
78	Consultee 2	3.10	The Cochrane review into teledermatology noted that the data included in the review were unlikely to be generally applicable. We note the specificity of 84.3% from the Cochrane review combined with the fact that that 6% of urgent suspected skin cancer referrals result in melanoma or SCC diagnosis [1], would suggest a 79% discharge rate in a real world service. There is no published or anecdotal evidence for a discharge rate that high in urgent suspected skin cancer telederm pathways in the UK. - We have worked on over 20 pathways that have seen over 120,000 NHS patients over 4 years and have never seen a teledermatology discharge rate this high. - It significantly outstrips the discharge rate of 40-60% which the BAD felt could be achieved when reviewing "2ww" referrals [2], though we note that the evidence they cite shows discharge rates between 36% and 46% for lesion only pathways, with the higher numbers being from teledermatology pathways that were majority non-lesion pathways or pathways excluding suspected skin cancer cases [3]. Moreover, given that 6% of urgent suspected skin cancer referrals result in melanoma or SCC diagnosis [1] and BCC incidence is 2.4x that of melanoma and SCC (2018-2020 incidence rate of BCC was 261 per 100k population compared to 26 for melanoma and 83 for SCC) [4], it is likely that at least 20% of urgent suspected skin cancer referrals would have a melanoma, SCC or BCC diagnosis and a 79% discharge rate would risk missing cancers. Combined with the fact that the lowest number needed to biopsy for melanoma of specialists in a meta-analysis reported is 5.85 [5], a discharge rate that high is not possible.	Thank you for your comment, which the committee has considered. An analysis by the EAG suggested that, of eligible lesions, automated use of DERM could approximately halve the number of referrals to a dermatologist within the urgent skin cancer pathway. The company's early modelling suggested that automated use of DERM could result in more lesions being correctly identified as non-cancer without a biopsy compared with teledermatology or face-to-face assessment. So fewer biopsies would be needed, and people would be correctly discharged from the service. The committee noted that a well-established teledermatology service could also reduce the number of referrals to face-to-face dermatologist appointments. It discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.

			Furthermore, as commented in section 3.9, the AI triage sits before teledermatologist review in the pathway, ensuring that the case volume is safely reduced ahead of teledermatologist review, at which point Trust dermatologists can discharge more patients. 1. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection 2. https://bad.org.uk/clarification-on-the-challenge-to-the-bad-2ww-letter-and-use-of-ai/ 3. https://onlinelibrary.wiley.com/doi/epdf/10.1111/bjd.17904 4. https://www.cancerdata.nhs.uk/getdataout/skin 5. https://doi.org/10.1016/j.jaad.2019.12.063	
79	Skin Analytics	1.3	While we agree the teledermatology data in the Cochrane review does not generalise to the current UK skin cancer pathways, it is not clear that the COVID 19 pandemic is relevant given that this generalisability issue was raised by the authors raised at the time of publication before the pandemic. Further, these limitations were raised with respect to the teledermatology data specifically and not with respect to the other analyses, including those of in person assessments. While urgent suspected skin cancer pathways have been changing with the rollout of teledermatology in more regions, the Technical Annex associated with the Darzi Report highlights that less than half of Urgent Suspected Cancer Referrals were managed via teledermatology [1]. In that context, it is not credible to suggest that pre-pandemic evidence should not be considered by the committee. Given the majority of patients see a dermatologist in person, this is still the more appropriate comparator. It is unlikely that Dermatologist in person sensitivity should have changed after the pandemic, given that a recent paper published in Nature suggested that Dermatologist sensitivity for melanoma was 74% which is at the lower end of previous studies [2]. While we value clinical expertise highly, it does not confer expertise at inferring clinical performance of clinicians across the UK. In the most recent public committee meeting, a specialist committee member suggested with no supporting evidence that teledermatology was	Thank you for your comment, which the committee has considered. Section 3.5 now notes that the committee agreed that there is no evidence to suggest that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identifying cancer lesions.

			likely more accurate at finding non-cancer lesions than reported in the Cochrane review based on teledermatology being more widely used. However other experts have suggested that as teledermatology has increased, less experienced dermatologists and non-dermatologists (e.g. Plastic Surgeons and GPs with Extended Responsibilities) are now doing teledermatology and therefore it is likely that the accuracy of teledermatology is lower than reported. If the committee do wish to consider teledermatology as a comparator, we would highlight that a meta-analysis published in Nature this year found a sensitivity of teledermatology for skin cancer of 84.2% [3], while the NHSE/Edge Health report meta-analysis found a teledermatology sensitivity for melanoma of 92.2% [4], both of which are lower than DERM [5]. Taken together it is not clear that the lack of clear evidence for or against teledermatology introduces a risk to patients when it is clear DERM performance is comparable to in person dermatologist assessment. 1. https://assets.publishing.service.gov.uk/media/66e1b517dd4e6b59f0cb2553/Independent-Investigation-of-the-National-Health-Service-in-England-Technical-Annex.pdf 2. https://doi.org/10.1038/s43856-024-00598-5 3. https://doi.org/10.1038/s41746-024-01103-x 4. https://www.england.nhs.uk/blog/exploring-the-future-artificial-intelligence-ai-and-dermatology/ 5. https://skin-analytics.com/ai-pathways/derm-performance/	
80	Skin Analytics	3.10	A Cochrane review showed 94.9% (95% CI 90.1-97.4%) sensitivity of teledermatology for skin cancer (N=717) [1]; a meta-analysis published in Nature this year showed 84.2% (95% CI 76.2–89.8) sensitivity of teledermatology for skin cancer; and the meta-analysis conducted by Edge Health commissioned by NHS England this year showed 92.2% (95% CI 87.3-97.0%) sensitivity for melanoma [3]. Please also see comments in section 1.3.	Thank you for your comment, which the committee has considered. See response to comment 79.

			1. https://doi.org/10.1002/14651858.CD013193 2. https://doi.org/10.1038/s41746-024-01103-x 3. https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf	
81	Skin Analytics	3.10	A specialist committee member raised this point in the public committee meeting and stated that the accuracy is likely higher, but without supporting evidence. As written this statement implies actual teledermatology accuracy is greater than reported due to more dermatologists doing it which is not necessarily the case and as such is misleading. There are a number of factors which may actually worsen teledermatology performance, such as fewer experienced dermatologists as they retire, lower levels of teledermatology experience amongst dermatologists taking it up more recently and non-dermatologists (e.g. Nurse Practitioners, GPwER) known to be conducting teledermatology reviews. Is the implication that teledermatology accuracy would surpass in person dermatologist sensitivity? If not, the existing literature for in person dermatologist performance is a more appropriate benchmark.	Thank you for your comment, which the committee has considered. See response to comment 79.
82	Skin Analytics	3.10	These studies are not outdated as there is considerable variation in the uptake of TD in the dermatology community and so represent the real world. The only notable difference between services at the point at which the Cochrane review was published and today is that a higher percentage of assessment is done via TD than it was previously. Our data show us that TD has a lower NPV than F2F assessment, so while it may be the case that services are more cost effective than previously due to the introduction of TD, it is also likely that the rate of false negatives today is higher than the Cochrane review would suggest.	Thank you for your comment, which the committee has considered. Section 3.12 of the guidance has been updated.
83	Skin Analytics	3.5	It should be noted that biases may exist amongst clinicians who overestimate their own performance. Especially when it comes to TD. Indeed There is a lack of evidence to show how many melanomas can be picked up by AI when this is applied to lesions who have been diagnosed by	Thank you for your comment, which the committee has considered. Section 3.5 now notes that the committee agreed that there is no evidence to suggest

			dermatologists as being benign. Given that both AI and dermatologists return false negatives, it is illogical to base assessment only on melanomas picked up by dermatologists checking lesions after AI screening.	that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identifying cancer lesions.
84	Birmingham AI and Digital Health Research Team, University of Birmingham	1.3	We agree that further evidence is required regarding the use of teledermatology with and without AI to identify the respective performance levels. This applies to patient outcomes but also cost-effectiveness.	Thank you for your comment, which the committee has considered.
85	Birmingham AI and Digital Health Research Team, University of Birmingham	1.3	We would like to encourage the NICE EVA committee to be more specific with regards to the level of evidence required for teledermatology alone versus with AI. The Edge health report written for the NHSE Outpatient Recovery and Transformation Programme has compiled data regarding teledermatology services. The negative predictive value for teledermatology is noted as 97.6%[95.5%-99.6%], f2f evaluation 98.0% [97.1%-98.9%]. Based on the literature the NPV for DERM is at least as good as both of these alternatives.	Thank you for your comment, which the committee has considered. The committee concluded that based on the data from Marsden et al. (2024) and Marsden et al. (2023), DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer pathway. Compared with face-to-face assessment, results suggest that DERM used in teledermatology services would discharge fewer non-cancer lesions from the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.

Theme: Comparison with face-to-face assessment

Comment number	Name and organisation	Section number	Comment	NICE response
86	Consultee 2	2.6	You correctly identify that there are two comparators. All your comments refer to effectiveness and cost-effectiveness relative to teledermatology. You should also comment on the effectiveness and cost-effectiveness relative to face-to-face consultation to be consistent with your original scope	Thank you for your comment, which the committee has considered. DERM 005 (Marsden et. al 2024) reported a sensitivity of 94.0% (95% confidence intervals [CI]: 84.7 to 98.1) for automated DERM in a real-world setting compared to 97.0% (95% CI: 88.7 to 99.5) for teledermatology. DERM 003 (Marsden et al. 2023) reported sensitivities for detecting cancer lesions of 96.0% (95% CI: 92.6 to 98.0) for automated DERM compared to 93.8% (95% CI: 90.0 to 96.3) for face-to-face dermatologist review. Neither DERM 003 nor DERM 005 did a test of equivalence, so it is not certain that DERM has equivalent sensitivity to teledermatology and face-to-face dermatologist review for detection of cancer lesions. But, the confidence intervals overlapped in both studies, and the committee agreed that there is no evidence to suggest that DERM is less sensitive than teledermatology or face-to-face dermatologist review for identification of cancer lesions. The committee concluded that further evidence should be generated on the sensitivity to detect cancer lesions of automated DERM used within a well-established teledermatology service compared with the sensitivity of a well-

				established teledermatology service alone, as well as the impact this has on dermatologist capacity. See section 3.5 of the guidance. Marsden et al. (2024) reported specificities for detection of cancer lesions of 73.3% (95% CI: 69.9 to 76.4) for automated DERM in a real-world setting and 71.9% (95% CI: 68.4 to 75.1) for teledermatology. Marsden et al. (2023) reported the specificity for cancer detection with automated DERM to be 45.0% (95% CI: 39.5 to 50.6), which was considerably lower than the 77.4% (95% CI: 72.4 to 81.8) for face-to-face dermatologist review. These results suggest that DERM used in teledermatology services may have similar discharge rates to teledermatology services alone for triaging non-cancer skin lesions referred to the urgent suspected skin cancer pathway. Compared with face-to-face assessment, results suggest that DERM used in teledermatology services would discharge fewer non-cancer lesions from the urgent suspected skin cancer pathway. The committee concluded that further evidence should be generated on the specificity to detect cancer lesions of DERM used within a well-established teledermatology service compared with the specificity of a well-established teledermatology service alone, as well as the capacity impact this has on dermatology pathways. See section 3.6 of the guidance.

87	Skin Analytics	2.6	Although mentioned here in parallel to teledermatology, the face-to-face dermatology assessment has been omitted from the subsequent discussion. While urgent suspected skin cancer pathways have been changing with the rollout of teledermatology in more regions, the Technical Annex associated with the Darzi Report highlights that less than half of Urgent Suspected Cancer Referrals were managed via teledermatology [1]. As a result it would be inappropriate to over index on teledermatology. 1. https://assets.publishing.service.gov.uk/media/66e1b517dd4e6b59f0cb2553/Independent-Investigation-of-the-National-Health-Service-in-England-Technical-Annex.pdf	Thank you for your comments which the committee considered. See response to comment 86.
88	Skin Analytics	3.5	There is no mention of face-to-face assessment in this section, despite it being logged as a comparator in section 2.6. As per the previous comments on sections 1.3 and 3.4, DERM has been shown to be at least as accurate as dermatologists in safely ruling out skin cancer. That there is insufficient evidence regarding DERM compared to teledermatology does not reasonably lead to a conclusion that there is insufficient evidence of DERM performance given the clear evidence regarding the comparator of in-person dermatologist assessments. For this to hold, there must be reasonable evidence to support that teledermatology sensitivity is likely to be higher than in-person dermatologist assessments.	Thank you for your comments which the committee considered. See response to comment 86.
89	Skin Analytics	3.11	We know from the Darzi Report, Technical Annex, that the Q1 2024/2025 data demonstrated that less than half of urgent suspected cancer referrals are managed via teledermatology [1]. We know a significant proportion of these teledermatology pathways will have in fact been deployments where DERM was also deployed. Regardless, rather than “many areas” this statement should read “the majority of areas still refer all…”. In this context, it is not credible that face-to-face dermatologist performance as a comparator has been omitted from the committee discussion. 1.	Thank you for your comments which the committee considered. See response to comment 86.

			https://assets.publishing.service.gov.uk/media/66e1b517dd4e6b59f0cb2553/Independent-Investigation-of-the-National-Health-Service-in-England-Technical-Annex.pdf	
90	NHS England	2.6	It is unclear why the focus of comparison for AI is TD alone. It suggests that the view of NICE is that only one of these two assessment methods should be used in tandem with F2F assessment. In reality, the performance and cost effectiveness of the AI should be compared to all existing initial assessment types which are deemed safe (i.e. TD and F2F). It is reasonable and desirable that a service should utilise a mixture of all three assessment methods as appropriate.	Thank you for your comments which the committee considered. See response to comment 86.
91	British Association of Dermatologists		In order for NICE to robustly review evidence which compares the cost benefits for a patient referred for a USC, it needs to extend the comparison between AI, teledermatology and face-to-face. This would allow more robust comparison of the current shift towards pushing more patients on the urgent suspected skin cancer pathway through teledermatology services.	Thank you for your comments which the committee considered. See response to comment 86.

Theme: Use of DERM in black and brown skin

Comment number	Name and organisation	Section number	Comment	NICE response
92	Healthcare professional 1		From the performance data shared with us circa 1000 patients with skin types V and VI have been assessed by DERM and of these all 19 cancers were detected by the AI (we believe this data was validated by the Unity Insights assessment). Given that less than 0.5% of UK skin cancers occur in black and Asian patients a project to gather enough cases to validate the AI would take many years. In the meantime we know that a third of melanomas and SCCs are picked up on routine pathways (where there are up to a 12 month wait) and that this disproportionately affects Black and Asian patients, leading to worse patient outcomes. Our preferred approach is to deploy the autonomous technology with safety nets in place to identify patients with darker skin types routing them for a teledermatology review (which does not negatively impact patient care or experience) whilst continuing to build valuable data in this patient group to reduce inequalities.	Thank you for your comment, which the committee has considered. Sections 3.8 and 3.9 have been updated to include this detail.
93	Consultee 2	3.7	Understanding more about the accuracy of diagnosis of skin cancer in non-white skin is rightly identified as important research priority. However, increased understanding is a general requirement for all methods of diagnosis. Further the nature of the research is very difficult because of the rarity of skin cancer in non-white skin. The noted prevalence in the studies that have been done, all substantially below 5%, means that the scale of studies required to fully investigate accuracy would be enormous - it could possibly only be addressed at a national level. This would be very time consuming and costly. It seems unreasonable to delay a potentially effective technology when research on accuracy of diagnosis in non-white skin is required generally. A more reasonable and pragmatic approach may be to acknowledge that research in non-white skin will take time, but in the mean time the AI technologies should only be used in Fitzpatrick skin types where there is already good evidence of accuracy	Thank you for your comment, which the committee has considered. Sections 3.8 and 3.9 have been updated to include this detail.
94	Consultee 2	3.13	This is acknowledged. However, this should not be an argument for the technology not being used in patients where it has been demonstrated to be effective and coat-effective with reasonable certainty	Thank you for your comment, which the committee has considered.

				Committee concluded that because the amount of data remains small, more data is needed on the accuracy of automated DERM in people with black or brown skin to be sure AI technologies are not incorrectly detecting (false positive) or missing (false negative) skin cancer. See section 3.9 in the guidance.
95	Skin Analytics	1.3	These statements risk misleading the audience and if presented this way in the committee discussion may have adversely influenced their decision making. It should be clarified to explain that this reflects the incidence of disease and population who present to skin cancer pathways. Less than 0.5% of skin cancer diagnosed in the UK is in Black and Asian patients [1] and 3.3% of urgent suspected skin cancer pathway patients are Black and Asian [2]. Accordingly, 3.5% of patients in DERM pathways are identified as having Fitzpatrick types 5 and 6 skin. Please see further detail in our response to section 3.7. 1. https://doi.org/10.1038/s41416-022-01718-5 2. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection	Thank you for your comment, which the committee has considered. Section 3.8 of the guidance has been updated to include this detail.
96	Skin Analytics	3.7	This section has not attempted to detail how standard of care is performing for patients with black and brown skin in skin cancer pathways. There is substantial national data available, much of which was detailed in Skin Analytics' Equality and Health Inequalities Impact Assessment (EHIA) which was submitted. This context is vital when trying to interpret the performance of the technologies and we want to ensure the committee has considered this when reaching their decision. Less than 0.5% of skin cancer diagnosed in the UK is in Black and Asian patients [1] and 3.3% of urgent suspected skin cancer pathway patients are Black and Asian [2]. Accordingly, 3.5% of patients in DERM pathways are identified as having Fitzpatrick types 5 and 6 skin.	Thank you for your comment, which the committee has considered. The Skin Analytics EHIA report was considered by the EAG. Section 3.8 of the guidance has been updated to include the other information provided in the comment. The committee concluded that because the amount of data remains small, more data is needed on the accuracy of automated DERM in people with black or brown skin to be sure AI technologies are not incorrectly detecting

			At least 1 in 3 melanoma and SCCs are not found in the urgent suspected skin cancer pathway [2] (N.B. more recent provisional data from the Rapid Cancer Registration Dataset suggests this may be as high as 48% [3]) and 441,000 patients are currently waiting to be seen on these other pathways [4]. These diagnostic delays disproportionately affect Black and Asian patients - 48.9% of Black patients and 60.6% of Asian patients have their skin cancer diagnosed on the urgent suspected skin cancer pathway compared to 67.3% of White patients [2,5] and delays in melanoma diagnosis of 2 weeks or more are linked with 20% reduced 5-year survival [6]. The conversion rate of urgent suspected skin cancer referrals to skin cancer diagnosis in White patients (6.6%) is over 16 times that of Black (0.4%) and Asian patients (0.4%) [2] and the proportion of melanoma diagnosed at early stage (1 or 2) for Black and Asian patients is 79% and 86% respectively, compared to 94% or more in other groups [7]. DERM has assessed nearly 2,000 patients with Fitzpatrick types 5 and 6 skin and has so far picked up all of the melanoma, SCC and BCC assessed in these groups. Acral and subungual lesions are routed to Trust teledermatology review in the DERM pathway, ensuring safety in these more clinically-concerning lesion body locations. The robust Skin Analytics post-market surveillance programme ensures continued understanding of performance in these patients. The guidance also notes that “people from Black, Black Caribbean, Black African and Asian ethnic groups are more likely to have a worse prognosis because lesions may be detected later” but make no effort to acknowledge that the later detection may be from lower clinician diagnostic accuracy or that the current disparity with standard of care is not acceptable. Excluding patients with black or brown skin from accessing DERM pathways risks exacerbating the already significant health inequalities that exist; whereas our proposed post market surveillance approach supports use on black and brown skin while closely monitoring performance over time.	(false positive) or missing (false negative) skin cancer. See section 3.9 in the guidance.

			Further, DERM has been shown to consistently reduce the number of dermatologist appointments in pathways it is deployed. These appointments can be used for patients on non urgent skin pathways where, as stated above, the majority of Black (51.1%) patients and a significant number of Asian (39.4%) are diagnosed with their melanoma or SCCs compared to 32.7% of White patients. This section does not sufficiently investigate the benefits as well as the risks for people with black or brown skin. 1. https://doi.org/10.1038/s41416-022-01718-5 2. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection 3. https://www.cancerdata.nhs.uk/covid-19/rcrd 4. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 5. Public Health England. Routes to diagnosis 2015 update: malignant melanoma | National Cancer Intelligence Network Short Report [Internet]. National Cancer Registration and Analysis Service; 2015 Available from: http://www.ncin.org.uk/view?rid=3121 6. Pacifico, M, Pearl, R, and Grover, R. The UK government two-week rule and its impact on melanoma prognosis: an evidence-based study. Ann R Coll Surg England. (2007) 89:609–15. doi: 10.1308/003588407X205459 7. (March 2023-March 2024) https://www.cancerdata.nhs.uk/covid-19/rcrd	
97	Skin Analytics	3.7	This statement could be misleading and needs context to be accurate. DERM has been deployed as per its intended use within skin cancer pathways across NHS England. The patients who have been assessed reflect the populations who present to these pathways and the incidence of disease. DERM has evaluated a similar proportion of patients of Fitzpatrick 5 and 6 to the proportion that present under the current standard of care. Please can this statement be amended to reflect this.	Thank you for your comment, which the committee has considered. Section 3.9 of the guidance has been updated to note that automated DERM has primarily been evaluated in people with white skin (Fitzpatrick skin types 1 to 3) because of the lower incidence of skin cancer in black and brown skin.

98	Skin Analytics	3.7	This statement is misleading. Please can this be corrected to avoid misinterpretation of selection bias. The 3% of lesions assessed by DERM with confirmed diagnoses in Fitpatrick skin types 5 and 6 closely matches what is seen in national data for urgent suspected skin cancer pathways (3.3%) [1]. Please can this statement be amended to provide the context above. 1. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection	Thank you for your comment, which the committee has considered. See response to comment 97.
99	Skin Analytics	1.3	The evidence submitted shows that DERM has identified all 27 melanoma, SCC and BCC assessed in these patients. Excluding patients with black or brown skin from accessing DERM pathways risks exacerbating the already significant health inequalities that exist; whereas our proposed post market surveillance approach supports use on black and brown skin while closely monitoring performance over time. Please see further detail in our comments on section 3.7.	Thank you for your comment, which the committee has considered. Section 3.9 of the guidance has been updated to include that recent company data on using automated DERM in people with brown or black skin (Fitzpatrick skin types 5 and 6) showed that no cancer lesions were missed. This suggested that automated DERM is as diagnostically accurate in people with black or brown skin as it is in people with white skin.
100	Skin Analytics	1.3	The guidance is not clear on how much more data is required or whether due consideration was given to the thorough monitoring with ongoing post-market surveillance, with more clinically-concerning acral and subungual lesions routed for Trust teledermatology assessment. While it is important to continue to monitor DERM performance in such subpopulations, we are concerned that the committee has not considered the impact on patients with black and brown skin. With only 0.5% of skin cancer diagnosed in Black and Asian patients [1], gathering statistically significant data will take significantly longer than the two years the EVA process for evidence collection allows, even assuming a significant proportion of the UK population is offered access to DERM. The current standard of care has resulted in more than 441,000 patients	Thank you for your comment, which the committee has considered. The final guidance recommends that DERM can be used within teledermatology services in the NHS during the evidence generation period as an option to assess and triage skin lesions in adults referred to the urgent suspected skin cancer pathway. It also notes that the potential risk of missed or delayed cancer diagnoses when using DERM during the evidence generation period by should be

			waiting for a routine dermatology appointment as of the end of July [2]. 1 in 3 melanoma and SCC are found on routine pathways, a fact that affects patients with black or brown skin more significantly [3-4]. DERM has assessed nearly 2,000 lesions in Fitzpatrick 5 and 6 patients (1,981), with all 27 melanoma, SCC and BCC identified. Comparator data is not available for in person or teledermatology diagnostic accuracy for patients with black and brown skin though there is plenty of data showing that standard of care achieves worse outcomes for these patients and so it is unlikely that standard of care achieves this level of accuracy. In summary, while we agree more data is required, it is not clear that the committee has presented or reviewed data to assess the balance of risk and benefit for patients with black or brown skin. Specifically it would be in the public interest to understand how the committee evaluated the risk to patients with black or brown skin of using DERM with the proposed controls and weighted them against the potential benefits of capacity released for routine dermatology appointments. Without clear guidance on what amount of data would be sufficient as well as an estimate of the timeline required to gather this data, it is not clear a thorough evaluation has been conducted with regard to improving health inequalities. Please see further detail in our response to section 3.7. 1. https://doi.org/10.1038/s41416-022-01718-5 2. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 3. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection 4. Public Health England. Routes to diagnosis 2015 update: malignant melanoma | National Cancer Intelligence Network Short Report [Internet].	mitigated by doing a healthcare professional review for people with black or brown skin.

			National Cancer Registration and Analysis Service; 2015 Available from: http://www.ncin.org.uk/view?rid=3121	
101	Skin Analytics	3.7	DERM has assessed nearly 2,000 lesions in Fitzpatrick 5 and 6 patients (1,981), with all 27 melanoma, SCC and BCC identified. Comparator data is not available for in person or teledermatology diagnostic accuracy for patients with black and brown skin though there is plenty of data showing that standard of care achieves worse outcomes for these patients and so it is unlikely that standard of care achieves this level of accuracy. While the data needs to be collected for both standard of care and the intervention, there is sufficient evidence for benefits to patients.	Thank you for your comment, which the committee has considered. See response to comment 94.
102	NHS England	1.1	In the few percent of instances where the individual with a suspicious lesion has Fitzpatrick skin type 5 or 6, NHSE's clear position when working with trusts is that images taken for assessment by DERM are also passed to the dermatology service for review in a dermatologist-led teledermatology clinic. This "dual-pathway" forgoes productivity gains in those few instances but in effect maintains the second read for this demographic; allowing for continued collection of outcome data regarding the tool's efficacy on these skin types, without potentially compromising the quality or convenience of care experienced by these patients.	Thank you for your comment, which the committee has considered. The final guidance notes that the potential risk of missed or delayed cancer diagnoses when using DERM during the evidence generation period by should be mitigated by doing a healthcare professional review for people with black or brown skin.
103	NHS England	1.3	As per previous comments, this is a position with which NHSE agree	Thank you for your comment, which the committee has considered.
104	NHS England	3.6	Please see previous comments regarding Fitzpatrick skin types V-VI	Thank you for your comment, which the committee has considered.
105	NHS England	3.7	The device's performance in releasing clinical capacity when used autonomously does not rely on its ability to detect malignancy but rather benign lesions. The ongoing data collection will contribute to the answer to this question but initially this is where a human clinical review is important. It should be noted here again that NHSE do not plan to move to an autonomous model for this group of patients until further evidence has been generated and service evaluation completed.	Thank you for your comment, which the committee has considered. The committee agreed with this approach. See response to comment 102.

106	Unity Insights and University of Surrey	3.6	Differences in incidence across the Fitzpatrick skin type subgroups were not significant and all confirmed cancers were correctly classified in the group representing skin types 5 and 6. Whilst encouraging, these results should be confirmed by further analyses due to the small percentage of Fitzpatrick skin types 5 and 6 cases (only 4.1% of assessed cases) in the dataset.	Thank you for your comment, which the committee has considered. See response to comment 99.
107	Birmingham AI and Digital Health Research Team, University of Birmingham	1.3	We agree that further data is required with respect to how AI technologies perform for darker skin tones. This is however extremely difficult evidence to generate given the prevalence of cancers in groups with darker skin tones. We would recommend that this evidence is generated in the post-deployment phase, and these groups are closely monitored (potentially with a second read for these groups). If the NICE EVA committee require this evidence to be generated before technologies are deployed, then specific recommendations around how this should be done would be helpful.	Thank you for your comment, which the committee has considered. See response to comment 102.

Theme: Measuring skin tone with skin spectrophotometry

Comment number	Name and organisation	Section number	Comment	NICE response
108	Birmingham AI and Digital Health Research Team, University of Birmingham	3.7	Although we agree with the fact that Fitzpatrick skin type is not the ideal way of measuring skin tone, it has become the standard metric for comparison. The recommendation of spectrophotometry is relevant however we recommend that this is a wider initiative potentially beyond the scope of the EVA. This change in practice will need to be sustainable, beyond the evidence generation phase.	Thank you for your comment, which the committee has considered. Clinical experts advised that studies should measure skin tone with spectrophotometry rather than using the Fitzpatrick scale because spectrophotometry is a more accurate way of measuring total melanin content in skin. See section 3.9 of the guidance.
109	Consultee 2	3.7	This would further add to the complexity of the research. If spectrophotometric categorisation of skin tone was required, this would exclude the possibility for investigation using real world research approaches because spectrophotometry is not done routinely. It would require a specially designed and commissioned study which would add to cost and delay.	Thank you for your comment, which the committee has considered. See response to comment 108.
110	Skin Analytics	3.7	While we appreciate the value of spectrophotometry as an ideal, work is still ongoing as to how best to apply it and it is not standard of care. It may be unethical to exclude patients from a technology with benefits outlined in earlier comments while spectrophotometry data is collected. As outlined earlier, DERM has assessed nearly 2,000 lesions in Fitzpatrick 5 and 6 patients (1,981), with all 27 melanoma, SCC and BCC identified. Meanwhile, over 441,000 patients are waiting to see a dermatologist on routine pathways [1], where 1 in 3 melanoma and SCC are found and this is known to affect patients with black and brown skin more significantly [2]. There is no reason why this data needs to be gathered in a research setting, but instead can be introduced as real world evidence monitoring for all pathways when an effective scalable spectrophotometer is available.	Thank you for your comment, which the committee has considered. See response to comment 108.

			1. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2024/09/RTT-statistical-press-notice-Jul24-PDF-386K-88372.pdf 2. https://www.cancerdata.nhs.uk/cwt_conversion_and_detection 3. Public Health England. Routes to diagnosis 2015 update: malignant melanoma | National Cancer Intelligence Network Short Report [Internet]. National Cancer Registration and Analysis Service; 2015 Available from: http://www.ncin.org.uk/view?rid=3121	
111	NHS England	3.7	NHSE welcome this feedback and would be happy to work with any services utilising the technology to implement this and therefore more accurately determine which patients should receive additional trust safety netting as previously outlined. It should however be noted that spectrophotometry is not widely available and so use of this technology would not be an appropriate condition for deployment.	Thank you for your comment, which the committee has considered. See response to comment 108.
112	Healthcare professional 1		To reduce inequalities we need to mitigate risk whilst continuing to collect data. Are teledermatology and face to face services also collecting this data? You will never get the level of data you need to validate performance using spectrophotometry (which is not widely available outside research settings and the parameters for which have not been agreed) or in a purely research setting (see comment in 1.5). Therefore inequalities for these patient groups will continue to be perpetuated.	Thank you for your comment, which the committee has considered. See response to comment 108.

Theme: Eligibility for assessment

Comment number	Name and organisation	Section number	Comment	NICE response
113	Skin Analytics	3.8	This is inaccurate and was clarified during the committee meetings. Patients who are not eligible for DERM assessment do not need face-to-face appointments and would be assessed by teledermatology where an appropriate next step could be determined in the majority of instances. For example, our most recent data shows that out of nearly 3,000 cases not eligible for DERM assessment but reviewed by teledermatology at Chelsea & Westminster, 90% avoided an urgent suspected cancer face-to-face appointment.	Thank you for your comments which the committee considered. This has been updated in section 3.10 of the guidance which states that the committee noted that a large proportion of skin lesions that are not eligible for DERM assessment can be assessed by teledermatology.
114	Skin Analytics	3.8	We are satisfied with these assumptions which are based on our experience and data gathered running these pathways for over 4 years and having seen over 120,000 patients. However again we would draw the committee's attention to the fact that DERM and teledermatology work in series, not parallel. So the 9% of patients ineligible for DERM assessment but eligible for teledermatology would be evaluated by teledermatology and would accrue the benefits of teledermatology, as evidenced in the previous comment.	Thank you for your comments which the committee considered. See response to comment 113.
115	Skin Analytics	1.3	There is an inaccuracy in section 3.8 that suggests all lesions not suitable for DERM assessment must undergo face-to-face assessment, which seems to underpin this recommendation. There are some patients and lesions that need to be assessed face-to-face that are identified by appointment booking teams using referral form criteria; those that are suitable for teledermatology and DERM assessment; and those that are suitable for teledermatology but not DERM assessment, e.g. lesions too large for the dermatoscope. DERM is used in pathways that are able to combine the benefits of AI assessment and teledermatology review where indicated either due to the AI classification or where the patient or lesion meet the stated exclusion criteria that exist within the Instructions For Use. These pathways have been set up across more that 20 sites in ways that ensure patients are able to	Thank you for your comments which the committee considered. The EAG reported that the proportion of lesions that were excluded from studies because of lesion characteristics ranged between 15.6% and 27.4%, when reported. Excluded lesion characteristics included the lesions being obscured by hair or lesions that were mucosal, acral or involving nails. The company's data from NHS services which are already using DERM (collected from April 2020 to November 2023) reported that approximately 25% of lesions on the urgent suspected skin cancer

			access high quality clinical assessment. DERM has been deployed in this way for more than 4.5 years in pathways that have seen >120,000 NHS patients. We challenge the need for additional data relating to this.	pathway were not eligible for DERM. The company's economic model assumed that fewer people were eligible for assessment by automated DERM than teledermatology (81% compared with 90%). The committee concluded that it was appropriate for further evidence to be generated on the proportion lesions that are eligible for assessment by DERM compared with teledermatology alone. See section 3.10 of the guidance.
116	Skin Analytics	3.8	It is unclear why this is considered an evidence gap. We have seen >120,000 NHS patients across over 20 pathways over 4 years and have shared data on the proportion eligible for DERM assessment and the proportion not eligible for DERM assessment but still suitable for teledermatology assessment. We would again ask the committee to consider how teledermatology alone is expected to meet the unmet need given the scale of dermatologist workforce shortages. See comments in section 1.3.	Thank you for your comments which the committee considered. See response to comment 115.
117	Healthcare professional 1		Of the lesions referred from primary care on the urgent suspected skin cancer pathway approximately 10-15% are not suitable for teledermatology and are booked face to face assessments (as part of a admin triaging process). The remaining 85-90% are booked for teledermatology clinic. Of these there is approximately 1% technical failure rate (usually due to wifi connectivity) and approx 24% of lesions are excluded from AI assessment due to lesion characteristics (eg on hair bearing, mucosal, acral or nails etc). The remaining 75% undergo AI +/- teledermatology assessment.	Thank you for your comments which the committee considered.
118	Healthcare professional 1		This is incorrect a proportion of lesions are ineligible for AI assessment but they still go on to have a teledermatology review. Only 6% of our teledermatology / AI patients need an urgent face to face clinic review. In total 10-15% of all urgent skin cancer referrals are not suitable for TD assessment (the majority of these are identified prior to the patient attending) this is irrespective of the AI technology.	Thank you for your comments which the committee considered. See response to comment 113.
119	Healthcare professional 2		The exclusion criteria are quite limiting – so reduced suitability /flexibility compared to a Consultant dermatologist	Thank you for your comments which the committee considered.

120	Consultee 2	3.8	There is good data on the proportion of referrals which are eligible for DERM in the many patients who have already been assessed with DERM in the NHS. Further data collection is not justifiable	Thank you for your comments which the committee considered. See response to comment 115.

Theme: Referral rates

Comment number	Name and organisation	Section number	Comment	NICE response
121	Consultee 2	3.9	The economic modelling does also illuminate whether referrals would be be reduced by DERM, relative to teledermatology, as well as providing estimates of cost-effectiveness. Although there is some uncertainty, the number of referrals is likely to be less with DERM than teledermatology	Thank you for your comments which the committee considered. The committee discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.
122	Consultee 2	3.9	The section highlights that modelling of the effect of DERM on referrals may be an acceptable alternative to empirical demonstration of reduced referrals. The recommendations for further research do not seem to acknowledge that such modelling could provide sufficient evidence base without requiring real world confirmation. The use of such "linked evidence" approaches has been extensively used by NICE in the past, so it is a surprise that it is not being allowed in this technology	Thank you for your comments which the committee considered. See response to comment 121.
123	Consultee 3	1.3	Capacity issues are certainly addressed by telederm review. Simple maths using previous NHSE data on rapid access derm referrals of 500,000 pa: 3x more patients can be seen in a telederm clinic than a face-to-face (f-2-f) with lower resources consumed. Diagnosing and removing two thirds of the referred cases means only one third need f-2-f assessment many of which may be direct to surgery further reducing first clinic assessments. Assuming ±15 patient's per f-2-f and including the teledermatology assessment of a clinic one can extrapolate to reduce demand by 12,000 clinics or the full time a capacity of 40+ dermatologists. (I have data demonstrating a 90% reduction in skin lesion referrals - with caveats)	Thank you for your comments which the committee considered.

			Services can be set up 'pre-referral' in the community setting by shifting personnel into that space, but the usual secondary and tertiary care ivory towers exist and barriers are constructed by associations and official institutions in a self-serving way rather than for patient's benefit or safety as is usually espoused.	
124	NHS England	3.9	NHSE agree with this statement	Thank you for your comments which the committee considered.
125	Unity Insights and University of Surrey	3.8	Reduction in unnecessary referrals: DERM discharge rate (between 31 to 54% across all three secondary care sites) suggested a significant number of low-risk lesions had been removed from the secondary care 2WW pathway, reducing the number of cases considered by trust teledermatology services and increasing case conversion rates (including basal cell carcinoma) to 12.6% following DERM assessment, compared with 8.2% from GP referral. In the community care setting, assessing pre-referral lesions/cases resulted in a higher DERM discharge rate (60.1%). The case conversion rate of 18.2% was similar to the secondary care post-referral model rate (12.6%), but at the expense of assessing greater numbers of lesions which would otherwise have been seen by GPs. Expedition of diagnosis/treatment: Expedition of diagnosis/treatment was difficult to assess directly due to gaps in the site data once case management decisions were made, as well as the lack of granularity in data provided by sites pre-deployment. Therefore, no conclusion could be made regarding earlier diagnosis and reduction in waiting times. Second Read: The second read (dermatologist review of cases flagged for discharge) decision making was risk averse, overturning a number of potential DERM discharges, which may lead to unnecessary appointments. The second read overturn rate was 36%, and of the overturned cases, 41% were subsequently discharged following virtual review by trust dermatologists. Across the evaluation, 7 cancers were caught among 754 cases flagged by the second read for additional trust review. There was no evidence of incorrect discharge by the second read, through patients returning within six months.	Thank you for your comments which the committee considered. The committee discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.

126	Healthcare professional 1		Our experience of DERM in our urgent suspected skin cancer pathway is that over 95% of patients avoid an urgent face to face appointment with a clinician. We have comparative data on teledermatology pre and post AI depoloyment. Approximately 6% of referrals require an urgent face to face clinician appointment whether seen by teledermatology alone or with AI so this has remained static. However following AI deployment the number of patients requiring routine follow up appointments reduced by 13% and the number of patients needing a skin biopsy reduced by 10%. Together with autonomous AI discharges and streamlining the pathway across our trust this generates an additional 380 Programmed Activities (PAs, 1PA = 4 hours of clinician time) worth of very much needed clinical activity reducing our reliance on costly insourcing and working list initiatives to reduce our routine waiting list.	Thank you for your comments which the committee considered. See response to comment 125.
127	Healthcare professional 1		3.9 We have already shown that use of AI reduces routine face to face appointments by 13% and minor operations by 10% compared with a highly efficient teledermatology service. The benefits over traditional face to face appointments are even greater.	Thank you for your comments which the committee considered. See response to comment 125.
128	Skin Analytics	1.3	In the DERM post-referral pathway, DERM assessment comes first and patients are discharged or routed to Trust teledermatology. The Trust teledermatologist can discharge patients at their review too. As the AI triage sits before teledermatologist review in the pathway, the reduction in case volume happens ahead of teledermatologist review. This is well evidenced as per our previous comment and we have seen across our deployments that have now seen 120,000 NHS patients that the case volume is consistently and meaningfully reduced by DERM. We reference again the criteria for an EVA which requires the balance of risk and benefit to NHS patients to show sufficient promise for ongoing use while additional data is collected. It is not reasonably concluded that the lack of comparator data on referral numbers with AI powered teledermatology and teledermatology introduces significant risks to NHS patients.	Thank you for your comment, which the committee has considered. In the assessment comparisons were made between DERM used within teledermatology services and teledermatology services alone, and DERM used within teledermatology services and face-to-face dermatology assessment alone. See section 2.7 of the guidance.

Theme: Primary care referral rates

Comment number	Name and organisation	Section number	Comment	NICE response
129	Birmingham AI and Digital Health Research Team, University of Birmingham	3.9	It is also important to note that the implementation of DERM can affect the likelihood of primary care practitioners to refer. There might be value in understanding how referral rates and referral behaviors change before and after deployment of DERM.	Thank you for your comment, which the committee has considered. The population in the scope is people who have been referred on the urgent suspected skin cancer pathway. Impacts on decision making that happen before referral to the urgent suspected skin cancer pathway are outside the scope of this assessment.
130	NHS England	1.3	More evidence is needed here which will be generated by an increased data collection in a real world, service evaluation setting. It is to be noted that <0.5% of malignancies in the UK are from people from black or Asian heritage and so large datasets will support reducing of health inequalities and a human clinical review may continue to be appropriate in a small subset of our population. Current deployment protocols are clear that those with a Fitzpatrick skin type V-VI who are assessed by AI should also routinely have their images reviewed in a trust teledermatology clinic. This dual pathway supports NICE's view that more data should be gathered before autonomous use is recommended for these skin types, whilst still allowing services to benefit safely from productivity gains of autonomous use on patients with skin types I-IV. The broader dataset of real-world usage of DERM now contains ~90,000 patients (~88,000 of which are from skin types 1-4 due to overall prevalence rates of suspected skin cancer in the UK).	Thank you for your comment, which the committee has considered. See response to comment 129.
131	NHS England	1.3	Referral data at trust level for the USSC in England is publicly available and this data will have been considered by the committee as part of its evidence review.	Thank you for your comment, which the committee has considered.

			Given that the device is deployed primarily in a post-referral model (i.e. a clinician in primary care has already deemed a referral appropriate) we do not expect there to be an unwarranted rise in referrals, and this has not been noted in the pilot sites.	See response to comment 129.
132	British Association of Dermatologists		There needs to be research to map patient behaviour with widening of AI use, e.g. awareness of 'quick-check' AI leading to increased primary care consultations to seek reassurance. This would likely lead to more 'uncertain diagnosis- advise refer' generated from an AI tool then referred from primary to secondary care – hence increasing service demand. We need to know if this meets unmet need and earlier cancer diagnosis or actually further congests Service without health gain.	Thank you for your comment, which the committee has considered. An evidence generation plan has been developed to provide further information on the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.

Theme: MoleAnalyzer pro

Comment number	Name and organisation	Section number	Comment	NICE response
133	Skin Analytics	1.3	Whilst this statement is accurate that DERM can determine that patients do not require further assessment, the intended use and regulatory clearance of Moleanalyzer pro do not allow this.	Thank you for your comment, which the committee has considered. The committee agreed that Moleanalyzer pro is out of scope for this assessment because it is not intended to be used for triaging lesions referred on the urgent suspected skin cancer pathway. See section 2.3 in the guidance.
134	NHS England	1.1	It remains the position of NHSE that this technology is out of scope for the early value assessment. The aim of this assessment when it was requested by the NHSE outpatient recovery and transformation programme was to synthesise all available evidence regarding AI's ability to triage benign lesions away from the urgent suspected skin cancer pathway without the need for dermatologist review. Moleanalyzer pro does not have regulatory approval for this use case and is not being used in this way.	Thank you for your comment, which the committee has considered. See response to comment 133.
135	NHS England	1.3	As per our previous comment, we do not believe that Moleanalyzer Pro should be considered in this EVA. It is intended as a diagnostic aid to be used by clinicians and is therefore not suitable for the stated use case.	Thank you for your comment, which the committee has considered. See response to comment 133.

Theme: Equalities

Comment number	Name and organisation	Section number	Comment	NICE response
136	Skin Analytics	3.13	Please see the comments in section 3.7 regarding how standard of care currently impacts patients with richer skin, as well as summarising the current evidence that DERM has assessed nearly 2,000 patients with Fitzpatrick 5 and 6 skin, so far catching all 27 melanoma, SCC and BCC identified. As per comments in sections 1.3 and 2.5, we know that delayed melanoma and SCC diagnosis due to referral on the wrong pathway disproportionately impacts older patients in standard of care. In our pathways, we have seen patients as old as 102 go through successfully. Age is often incorrectly used as digital exclusion but our pathways have healthcare professionals capturing patient information and this should not be a criterion for type of assessment on its own but rather other factors such as mobility, ability to consent and so forth. Most of the equality considerations referenced in this paragraph pertain as much if not more so to teledermatology than they do to AI technologies and teledermatology has been deemed appropriate for wide scale deployment across the NHS on the balance of risk and benefit to patients.	Thank you for your comment, which the committee has considered. Section 3.15 of the guidance has been updated to remove older age as an equality issue.
137	Healthcare professional 1		There is some confusion here, it should read 'Teledermatology may not be suitable for people with more than 3 lesions.' We have many elderly patients attending our teledermatology service and have not noted a worse outcome for this group or worse patient acceptability. The arguament for 'whole body skin examinations' applies to teledermatology services and is not specific to AI technology.	Thank you for your comment, which the committee has considered. Section 3.15 of the guidance has been updated to remove older age as an equality issue. Section 3.4 of the guidance highlights the benefit of face-to-face dermatologist assessment enabling full body assessment, while acknowledging that capacity issues in the system meant that full body assessments are currently not practical.

Theme: Cost-effectiveness

Comment number	Name and organisation	Section number	Comment	NICE response
138	Consultee 2	3.10	It should be noted in the guidance that estimate of specificity used by DERM in the cost-effectiveness model is again conservative, with Marsden et al 2024 suggesting that the specificity of DERM is considerably higher than 42%, and definitely likely to be higher than teledermatology. The finding that DERM is cost-effective even with cautious assumptions about specificity should be noted in the guidance	Thank you for your comment, which the committee has considered. The company model assumed that automated DERM has a specificity of 42% based on real-world performance data. Specificity of DERM with a healthcare professional review would be lower than the specificity of automated DERM. So, the cost effectiveness of DERM used within a teledermatology service with and without healthcare professional review compared with teledermatology alone is uncertain. See sections 3.12 of the guidance.
139	Consultee 2	3.8	This is the first time an economic model has been criticised for being cautious in a way which disadvantages the technology! The estimates on the model data were based on real world data and are hence one of the parameters where there is relatively little uncertainty. Importantly, in the Skin Analytics economic model, the difference in eligibility is fully represented in the model, and the generally good estimates of cost-effectiveness exist even in the face of increased rates of ineligibility of DERM.	Thank you for your comment, which the committee has considered.
140	NHS England	1.3	The data gathered by the independent evaluation commissioned by NHSE are clear that there are economic efficiencies associated with the use of DERM compared to existing initial review mechanisms	Thank you for your comment, which the committee has considered. The committee concluded that the cost effectiveness of DERM used within a teledermatology service with and without healthcare professional review compared with teledermatology alone is uncertain.

141	NHS England	3.10	If clinical capacity were not limited and it was achievable for services across the country to collectively hire additional dermatologists to undertake the extra TD clinics required to reduce waiting lists, then this would be a reasonable concern to have. However, given that dermatology clinical capacity is stretched and that the increase in workforce at one trust is usually accompanied by a corresponding drop at another due to staff moving, we do not require that AI is cheaper in order for its introduction to be sensible. All we require is that it is priced at a level which is worthwhile for the clinical benefits of meeting a currently unmet need. Having said this, economic modelling from Edge Health is very positive regarding the cost effectiveness of autonomous AI vs human led assessment.	Thank you for your comment, which the committee has considered. See response to comment 140.
142	NHS England	3.11	The successful role out of Community Diagnostic Centres, https://www.nuffieldtrust.org.uk/resource/diagnostic-test-waiting-times, has provided a great opportunity to develop sites for image taking for TD. The costs are similar for TD and AI but autonomous AI comes with release of clinical capacity once embedded with assurance.	Thank you for your comment, which the committee has considered. The committee acknowledges there is an unmet need and section 3.1 of the guidance notes that AI technologies used within a teledermatology service could potentially increase staff capacity to help address the unmet need.
143	British Association of Dermatologists		The authors should be aware that this initiative is at risk of promoting an additional step in the routine pathway of patient care - use of AI in all referrals- with the additional costs entailed across the national Service. Whilst also arguing that we have currently too little evidence to argue for safe removal of the 'second-read', rapid adoption of AI model as described does risk the 'new norm' pathway to include AI - yes, potentially with service benefits, but with significant risks of increased Service costs without careful cost-benefit analysis.	Thank you for your comment, which the committee has considered. An evidence generation plan has been developed to provide further information on the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.
144	British Association of Dermatologists		The implications to our patients and total costs to the service of removal of the 'second-read' in AI-assisted triage needs substantially more evidence before consideration of adoption should occur. Evidence is required to	Thank you for your comment, which the committee has considered.

			address: - Accuracy of benign as well as malignant lesion diagnoses; - Are products capable of clear definition of type of benign diagnosis? - Would our population simply accept 'benign' or re-engage with HCPs to further probe type of lesion? - Cost of missed diagnoses to the NHS, etc.	The committee noted that: - Sensitivity for detecting cancer lesions is as important as sensitivity for detecting non-cancer lesions. This is because a test with a high sensitivity for cancer lesions will have a low number of false-negative results, that is, missed cancer lesions (section 3.5). - DERM's ability to classify non-cancer lesions is limited to 6 types of benign lesions and 2 types of pre-cancer lesions. The technology is not indicated to give a diagnosis of other types of benign lesion (section 3.7). - More evidence should be generated to understand the impact of using DERM on clinical capacity for both urgent and routine dermatologist services in the pathway (section 3.7). - DERM would not miss more cancers than a teledermatologist or a face-to-face dermatologist review of individual lesions (sections 3.5 and 3.16)
145	Consultee 4	3.10	Where is the analysis between cost effectiveness and delays due to lack of dermalogical staff and susequent treatment costs through lack of early identification or diagnosis?	Thank you for your comment, which the committee has considered. The EAG noted that in future modelling it would be important to consider how increases in staff capacity could be captured in the model, to meaningfully quantify the impact of reducing demand on dermatology services. See section 3.14 of the guidance.
146	Healthcare professional 1		3.10 Cost effectiveness of the AI over and above a highly efficient teledermatology service has been established at our trust. A business case has been approved and contracts signed for deployment as business as usual.	Thank you for your comment, which the committee has considered.

Theme: Future modelling considerations

Comment number	Name and organisation	Section number	Comment	NICE response
147	Birmingham AI and Digital Health Research Team, University of Birmingham	3.10	We wholeheartedly agree that there will be additional costs. One of the costs not mentioned here is also the cost of monitoring the AI system. This applies to the implementaiton phase but also during routine monitoring. There is a need for recurring local validation of all AIaMD as discussed here: https://www.nature.com/articles/s41591-023-02540-z.	Thank you for your comments which the committee considered. The committee acknowledged these strategies and information has been included in section 3.16 of the guidance.
148	Consultee 2	3.12	The additional complexity that would be required in order to do this is prohibitive. A whole dermatology service would need to be modelled in order to do this and would need to consider all the conditions treated by dermatologists, particularly the chronic skin conditions like eczema and psoriasis, whose care has been noted to be compromised by the amount of time needed to be invested in seeing suspected skin cancers. Estimates of time saved by reduced referrals can be done more simply, but have already been done in for instance in the study Marsden et al 2024, so we assume that this is not what is being suggested in this comment	Thank you for your comments which the committee considered. The committee discussed uncertainties on how using DERM in practice with or without healthcare professional review would impact capacity in the pathway compared with a well-established teledermatology service. The committee concluded that more evidence should be generated to understand the impact of using DERM with or without healthcare professional review on clinical capacity. See section 3.11 of the guidance.
149	NHS England	3.12	It is anticipated that there will be a release of clinical capacity and therefore a reduction in need of the work force growth that would be needed in line with demand growth with time. The assessment capacity which is added to services by the use of AI is in lieu of additional dermatologist workforce which is needed by services but is not currently available, and which would be relatively expensive and take time to recruit and train. Costs associated with the taking of images should be captured in the model,	Thank you for your comments which the committee considered. See response to comment 148.

			but it should be understood that these costs are a relatively cost effective method of meeting additional patient need.	
150	British Association of Dermatologists		There is a need to ensure that costing of future modelling of an AI-assisted pathway ensures patients with dark skin types are managed with an equal level of safety and accuracy as those with lighter skin types (where most evidence has been derived, to date). Therefore, the NICE EVA should consider the additional costs and ensure an equitable service for those patients where AI currently excludes (up to 25% of referred cases).	Thank you for your comments which the committee considered.
151	Healthcare professional 1		Also important to account for the fact that there are workforce shortages and that in reality the costs for seeing patients will therefore be escalated through waiting list initiatives and private insourcing which divert money away from departments to improve existing pathways.	Thank you for your comments which the committee considered.

Theme: Autonomous DERM

Comment number	Name and organisation	Section number	Comment	NICE response
152	Consultee 3	1.3	1.3 Questioning accuracy after granting Class II UKCA is closing the door of the horses bolted. This leads the industry to question the validity of the MHRA decision making that took everyone by surprise as it is an endorsement that the product is an independent clinical decision tool not a support tool. Then again: David Oliver: Our health regulators are in a crisis of competence and credibility | The BMJ If the safety of independent AI assessment is being questioned by this panel the legitimacy of the MHRA decision needs to be questioned as well. AI in isolation not as a teledermatology pathway is extremely limited. A lot more data has been generated in other tumour groups e.g. breast where AI acts as a second opinion and data is available in dermatology as well discussing the variability of sensitivity and specificity for AI versus clinical (although a lot of this is questionable as is the technician skill of clinician is usually not included) Using AI in conjunction with a clinician for safety should be an imperative until sufficient data is generated to prove that it can be used in isolation and that will take years.	Thank you for your comments which the committee considered. The committee acknowledged these strategies and information has been included in section 3.16 of the guidance.
153	NHS England	3.1	The autonomous use of AI in the urgent skin cancer pathway has the potential to release clinical capacity to be redirected to patients who require face to face consultations and ultimately to support elective recovery. We agree that uncontrolled adoption of these technologies without national oversight, informed consent, data collection and feedback is not appropriate.	Thank you for your comments which the committee considered. See response to comment 152.
154	NHS England	1.3	We agree that this is fundamental. These datsets exist and may be used for further evaluation.	Thank you for your comments which the committee considered.

Theme: Editorial

Comment number	Name and organisation	Section number	Comment	NICE response
155	Birmingham AI and Digital Health Research Team, University of Birmingham	1	The main concern we have regarding these recommendations, is that they are grouping together the different use cases/ positioning of AI in the pathway. A post-referral pathway will look very different to a primary care model with respect to cost-effectiveness, reduction in workload/referrals, and patient outcomes. We believe that this should explicitly addressed in the recommendations. Based on the previous point, we would agree that more research is needed for DERM in the primary care model, however this might look different for the post-referral pathway	Thank you for your comments which the committee considered. The positioning of DERM has been clarified in section 1.
156	Birmingham AI and Digital Health Research Team, University of Birmingham	3.9	Although we agree with the overall summary, we believe that there needs to be more explicit differentiation between the post-referral pathway and primary care pathway. With a post-referral pathway, the lower specificity of DERM compared to dermatologists alone is not the main issue. The latest version of DERM has a specificity of approximately 70%. Although this is lower than dermatologists specificity (85.7%), this will not increase the workload of healthcare staff. Without AI, 100% of referred cases will need to be seen by staff. With DERM, a significantly lower proportion would need dermatologist review. This can't be said about the primary care pathway and so we believe some of the recommendations should be tailored to the different pathways.	Thank you for your comments which the committee considered. The positioning of DERM has been clarified in section 1.
157	Consultee 2	3.8	The language is a little biased. If the percentage ineligibility rates are considered large, it would be more even handed to say that the ineligibility rates are high in both DERM and teledermatology. There is good data quantifying that ineligibility is greater for DERM (NICE have chosen not to include it for some reason), but that excess ineligibility with DERM is not "large", so again the language you are using is biased.	Thank you for your comments which the committee considered. Section 3.10 has been updated to state that a large proportion of skin lesions that are not eligible for DERM assessment can be assessed by teledermatology.

158	Consultee 2	3.9	This in agreement with the Skin Analytics model, which allows prediction of impact as well as cost-effectiveness. This should be mentioned in the guidance	Thank you for your comments which the committee considered.
159	Consultee 2	3.12	The guidance should note that the EAG conceptual model is very close to the model structure employed in the Skin Analytics economic model. This adds to the credibility of the cost-effectiveness findings produced by the Skin Analytics economic model	Thank you for your comments which the committee considered.
160	Skin Analytics	3.11	This is factually inaccurate, please can it be removed. DERM is an approved medical device that has been authorised to be deployed in line with its intended use. DERM is indicated for use on dermoscopic images of cutaneous lesions where there is a suspicion of skin cancer in patients aged 18 years or over except where specific exclusions apply. There is no limitation on the position in the pathway that it can be deployed.	Thank you for your comments which the committee considered. This has been clarified in section 3.13 of the guidance which now states: For DERM to be used in the post-referral pathway (that is, after a primary care referral), a local teledermatology service is needed.
161	Skin Analytics	3.11	This is factually inaccurate, please can it be corrected. Medical photographers are not specifically required as anyone can be trained to capture suitable images. The pathways where DERM is deployed commonly use healthcare assistants or nurses to capture images. Skin Analytics has worked with medical photographers to create training resources for image capturers, as well as providing on-site training.	Thank you for your comments which the committee considered. This has been updated in section 3.13 of the guidance which now states: An accurate DERM assessment relies on staff taking high quality medical photographs of the suspicious lesion. The staff are typically healthcare assistants, nurses or medical photographers who are trained to capture suitable images.
162	Skin Analytics	3.11	This is accounted for in the health economic model submitted, as was raised to be corrected at the last committee meeting.	Thank you for your comments which the committee considered.
163	NHS England	2.5	Please correct this as it is inaccurate and suggests that the NHS plan introduced TD for skin conditions which is incorrect. A virtual suspected skin cancer pathway was introduced following the publication of https://www.england.nhs.uk/wp-content/uploads/2022/04/B0829-suspected-skin-cancer-two-week-wait-pathway-optimisation-guidance.pdf in 2022. This new pathway provided the opportunity for images (including	Thank you for your comments which the committee considered. This section has been updated to the following: In 2022, the teledermatology pathway was introduced to support early

			dermoscopic images) to replace a F2F interaction on the urgent suspected skin cancer pathway. The use of teledermatology in referral optimisation for other skin conditions was recommended in https://www.england.nhs.uk/long-read/referral-optimisation-for-people-with-skin-conditions/ also in 2022	diagnosis of skin cancer. It provided the opportunity for images to replace face-to-face appointments for people referred to the urgent suspected skin cancer pathway. See section 2.6 of the guidance.
164	NHS England	3.11	This is factually incorrect. Whilst developing the TD roadmap NHSE worked closely with the Institute of Medical Illustrators to develop resources to upskill health care workers to take necessary images rather than recruiting additional trained medical photographers. In many places systems are now well established with Band 2 health care support workers trained to take images.	Thank you for your comments which the committee considered. Please see the response to comment 161.
165	Cancer Research UK		On the language: We support the advice for studies to measure skin tone with spectrophotometry. However, use of language such as “black and brown skin” is out of date and perhaps inappropriate. We propose NICE considers different language in the final publication for example a “range of skin tones”, whilst recommending measuring this range of skin tones via spectroscopy.	Thank you for your comments which the committee considered. The language used in the guidance follows the NICE writing style guide when referring to skin colour.
166	British Association of Dermatologists		The NICE EVA needs to recognize the risk of assuming the intended use case described - Hospital Level triage of suspicious lesions - could be extrapolated to use in a Primary Care setting without detailed evaluation of the safety and efficacy of its use as a decision support tool for Primary Care clinicians.	Thank you for your comments which the committee considered. The positioning of DERM has been clarified in section 1.
167	Healthcare professional 1		3.11 This is incorrect as per the teledermatology roadmap images can also be captured by HCAs and nurses using smartphones/tablets with dermatoscope attachments. In addition it is our understanding that the AI technology has been deployed in primary care in a number of locations.	Thank you for your comments which the committee considered. Please see the response to comment 161

Theme: General comments

Comment number	Name and organisation	Section number	Comment	NICE response
168	Consultee 1		What is the threshold for accepting AI in the first instance? What is the benchmark of confidence required for scanning decisions? Neither of these fundamental levels have been set or referred to establish a level to work from.	Thank you for your comment, which the committee has considered. An evidence generation plan has been developed to provide further information on the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.
169	Consultee 3	1.1	1.1 Why is the NHS fixated on these 2 systems in isolation? There are several systems out there some actually addressing some of the concerns/questions raised. Mole analyser is limited to pigmented lesions while rapid access referrals include all possible skin cancers - do you have any idea/data on the percentage of pigmented versus other lesions presenting in rapid access clinics as that would severely limit its value.	Thank you for your comment, which the committee has considered. The committee agreed that Moleanalyzer pro is out of scope for this assessment because it is not intended to be used for triaging lesions referred on the urgent suspected skin cancer pathway. See section 2.3 in the guidance.
170	Consultee 3	1.2	1.2 Moleanalyzer is developed by a multinational with deep pockets while DERM has received substantial support and millions in funding from NHSX/NHSE/DOH cancer collaboratives giving an unfair advantage to other start ups. Being under written by the NHS in this way is an endorsement and thus a commercial benefit as well. Full disclosure (and conflict of interest) of funding should be declared and the NHS endorsement as a diagnostic tool scrutinised by relevant bodies.	Thank you for your comment, which the committee has considered. See response to comment 169.
171	Consultee 3	1.3	The limitations of all systems as a generalisation rather than these 2 need to be identified and noted. Subungual lesions, interdigital, periocular sites, mucosal, palmar planter and darker pigmented skin may all impact but there is insufficient information on these. Alternative systems not solely relying on macro photography/dermoscopy should be explored.	Thank you for your comment, which the committee has considered.

			With teledermatology and AI failing certain sites/populations the benefit is reducing overall blocks to access face-to-face assessments by removing non-essential cases. Even clarity on what is being measured should be reviewed. Are BCCs included in the cancer identification data or not? Grossly dysplastic naevi and even in situ melanoma is not considered cancer but is still dealt with in the same way often with two-stage surgery (WLE margins) but is not cancer. Using MDT data excludes all of the above groups and no detailed studies exist on this.	
172	NHS England		Regarding the question: "Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?" This EVA provides summaries of the clinical and cost effectiveness analysis when comparing TD with autonomous AIaMD. A useful analysis would be using F2F as a comparator as this is a common real world experience. Further to this, NICE did not consider cost and clinical time savings resulting from reduced biopsies, as well as patient benefits including QALYs. There is a fundamental question that should also be asked as to the benefit of use of AIaMD being used as a triage tool to exclude the benign as all non-benign skin lesions are reviewed by a clinician The report suggests that there may be overdiagnoses of SCC and BCC, while this would not be the case as all suspicious lesions would be reviewed by a dermatologist. Additionally, NICE put forward as issues poor sensitivity for benign lesions as well as lower specificity than dermatologists, without accounting for the fact that currently all these lesions would be seen f2f, therefore any reduction in unnecessary reviews is a benefit	Thank you for your comment, which the committee has considered. The final guidance recommends that DERM can be used within teledermatology services in the NHS during the evidence generation period as an option to assess and triage skin lesions in adults referred to the urgent suspected skin cancer pathway.

			Evidence is quoted by NICE suggesting that AIaMD might lead to fewer missed cancers than current practice and a faster diagnosis, treatment and compliance with cancer outcomes for patients with malignant lesions. There are also more rapid discharges on the urgent suspected pathway which lead to cost savings. NICE has concluded that the tool does not show sufficient promise to be conditionally recommended but this tool is now being used in clinical practice. We suggest that there is sufficient promise for this tool to be used within the clinical parameters and guardrails of national commissioning including informed consent and a human clinical review for a period of time that ensures mandatory data capture and assurance of local expertise in embedding the tool. Data gaps on the comparison with TD are quoted as the main issue, without accounting for evidence of greater accuracy of DERM compared to teledermatology, meaning that AIaMD could be safer for patients than TD, as well as address critical workforce shortages.	
173	NHS England	2.1	It should be noted that option for benign diagnosis exists but many services have opted not to use this to date, which has left patients with limited information regarding their condition. Going forward, NHSE will be encouraging the use of this function to aid primary care and enhance patient experience.	Thank you for your comment, which the committee has considered.
174	NHS England	3.5	Diagnostic accuracy is not the measure of success for this device as we do not ask it to accurately diagnose malignancies (it passes all suspected malignancies onward for human review). Instead, its role is to triage away those lesions which it assesses as benign with the highest confidence. The sole use of sensitivity as a KPI is a weakness in NICE's evaluation of this technology.	Thank you for your comment, which the committee has considered. See response to comment 172.
175	Unity Insights and University of Surrey		Note that these comments are a collection of comments from the Technology Specific Evaluation Team comprised of the University of Surrey and Unity Insights who were the independent evaluators of DERM, commissioned by	Thank you for your comment, which the committee has considered.

			DHSC and NHSE, as part of the 3 year AI in Health and Care Award with Skin Analytics.	
176	Consultee 5		Disappointed that action is to be delayed. I lost my partner to melanoma; we need to do more to detect the disease early. The NHS is not in an ideal situation, there are insufficient staff. We need to grasp the opportunities that AI offers. Surely this too is in line with our new government's stated objectives to focus on early detection, through hubs in the community (similar to blood pressure checks?).	Thank you for your comment, which the committee has considered.
177	Cancer Research UK		Other evidence to consider: There is a white paper from Edge Health about DERM published in July 2024 and commissioned by NHS England which isn't taken into account in the consultation, yet may be of interest. It is a literature review of 27 studies (many of which are older than the studies included in the NICE consultation) - https://www.edgehealth.co.uk/news-insights/ai-in-dermatology-white-paper/.	Thank you for your comment, which the committee has considered.
178	British Association of Dermatologists		We have received a specific member comment on the lack of definition of what 'sufficient' evidence should be to mandate such a significant change in practice as well as a need to better describe the use cases in which we are looking for the AI products to operate, with significant risks of considering all products can operate across the different stages of a patient care pathway.	Thank you for your comment, which the committee has considered. An evidence generation plan has been developed to provide further information on the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.
179	British Association of Dermatologists		The BAD accepts the narrow focus of this EVA that the referenced products sit currently with the most evidence. In a fast-moving research field with multiple products, we have concerns that the EVA may preclude opportunity for rigorous assessment of other products entering the market within the EVA 3-year assessment process.	Thank you for your comment, which the committee has considered.
180	British Association of Dermatologists		We are keen to support evidence-based evaluation of AI across many potential use cases in Dermatology (secondary and primary care), but do have concerns that this EVA may bias and even hinder future research where incorrect assumptions may be made that the case for use in Dermatology is	Thank you for your comment, which the committee has considered.

			already proven, despite the single use case described and recognized deficiencies in the evidence base to date.	
181	British Association of Dermatologists		This is very high-level. Shouldn't the document give practical guidelines, e.g., the volume and types of cases etc, essential features of testing protocols, etc? Advise inclusion of clearer description of the type of AI tools being in scope for this EVA.	Thank you for your comment, which the committee has considered. Please see the published scope for further details on the AI technologies being considered for this assessment.
182	British Association of Dermatologists		The recent NHSE focus on rolling out teledermatology across urgent suspected skin cancer pathways has not been as robustly subject to cost benefit analysis and is still a relatively new and emerging concept for many Trusts compared to face-to-face '2WW' review.	Thank you for your comment, which the committee has considered.
183	British Association of Dermatologists		There is current work to develop an NHS-owned data pipeline of images and associated clinical DICOM metadata that could provide the critical resource for vendors to test their AI products for safe and effective use in the patient population managed in the NHS in the UK. This current report makes no suggestion of the value of such a company-agnostic approach to evidence-based assessment of AI products in this field.	Thank you for your comment, which the committee has considered.
184	Consultee 4	3.7	Whilst I understand the need for inclusivity of all skin types, it seems to me that in the UK the majority skin type is white and implementation of AI should not be held back awaiting futher information on black or brown skin.	Thank you for your comment, which the committee has considered. The committee concluded that the amount of data remains small. So, more data is needed on the performance of automated DERM in people with black or brown skin to be sure AI technologies are not incorrectly detecting (false positive) or missing (false negative) skin cancer. See section 3.9 in the guidance.
185	Consultee 4	3.9	So what needs to be done to allay this objection?	Thank you for your comment, which the committee has considered. An evidence generation plan has been developed to provide further information on

				the prioritised evidence gaps and outcomes, ongoing studies and potential real-world data sources. It includes how the evidence gaps could be resolved through real-world evidence studies.
186	Consultee 4	3.5	The committee should be more forthcoming on how much more research is required on each of the technologies and/or advise thresholds which need to be met.	Thank you for your comment, which the committee has considered. See response to comment 185.
187	Healthcare professional 1		1.3 What volume of data on DERM performance is needed? We have been using the DERM AI technology at our Trust in real world urgent suspected skin cancer pathways for over 2 years and assessed over 12,000 patients. In wider deployments over 100,000 NHS patients have been assessed with final outcome data for over 50,000 lesions. This is clinically relevant real world data and far surpasses the volume of data in the Cochrane review of Teledermatology (5,500 patients) which was deemed adequate to recommend national roll out as part of the NHS plan and outpatient recovery. Because of the way the technology has been applied it has essentially been assessing AI versus teledermatology.	Thank you for your comment, which the committee has considered. See response to comment 185.
188	Healthcare professional 2		Regarding the EDGE report and Negative Predictive Value - It is quite an expensive system to just be used to reassure the clinically benign lesions- lot of admin input, lot of patient time and attendance at appointment, - charge per use will mount up – no bulk buy discount available.	Thank you for your comment, which the committee has considered.
189	Healthcare professional 2		Skin Analytics are benefitting from the sense checking of their "private" company machine by being allowed access to the highly skilled validated NHS histopathology reports to train their algorithm. This data is a precious resource – have we under sold its value to a commercial company.	Thank you for your comment, which the committee has considered.
190	Healthcare professional 2		Running costs of the initial outlay – circa 160K?- We have to ask if the 2ww referral should have been made for the characteristic SEB K, and if it is not about who reassured the SEB K patient (AI or telederm or face to face but rather how do we stop the SEB K getting into the 2ww system at all) AI supported GP guidance to only refer lesions that meet certain criteria, as the bar is very low for a 2ww referral – Government target that detection rates for cancer referred as 2ww should be less than 2% - currently our conversions	Thank you for your comment, which the committee has considered.

			rate is around 6% - utilising face to face 2ww appointments The fact that it DERM AI can discharge 25-30% as benign may seem very attractive but it does beg the question , are the necessary thresholds being met in order to justify a 2ww referral. My personal experience from clinic this morning is that with workforce diversification to shore up primary care we are seeing an inexorable rise in "suspected skin cancer" referrals – These may not meet the quality or clinical index of suspicion to truly justify a 2ww referral- Very short duration of lesion Limited attempt at history Limited diagnostic clarity for benign lesions Partial treatment and fear of overlooking something so rapid resort to send 2ww – In GM there is a plan to work towards a single point of referral with referral AI led management to help guide GP and GP associates so that they can work within the bounds of Effective Use Resources and limit the referral of the benign lesions. GP education – or GP champion for skin cancer may be a more cost effective than AI. Less patient contacts- current derm AI pathway multi step pathway challenging for polder patients who would prefer single review and clear outcome. This is also being pursued by GM cancer with education available – but regrettably it is the ken dermatology aware GP who tend to come along on a Saturday PM for additional experience- so we educate those that least need education. Our Single point of referral may allow us to look at patterns of poor referral /incomplete form/history/criteria not clear for high index of suspicion (i.e. lesion has been there for years)	
191	Healthcare professional 2		I welcome the idea of support with management of 2ww demand – but wonder if AI at this point is the answer as many of the cases that DERM reassure are more than likely the cases that perhaps should never have been referred and turning the tap off at source might be much more cost effective that mopping up the overrunning bath. These are my personal views and do not represent the Trust.	Thank you for your comment, which the committee has considered.