Bias in Biomedical Datasets: Underrepresentation in Dermatologic AI

Group 2

MEDSCIEN9507

Dr. Wood

March 6, 2026

Interdisciplinary Medical Sciences

Schulich School of Medicine and Dentistry

Western University

Assignment Acknowledgement

To maintain transparency and uphold academic integrity, we acknowledge that we used ChatGPT, to assist with editing and reviewing our work for grammar and clarity.

Author Contributions

Student Name	Student Number	Contributions
Olayori Olasode	251215427	Contributed equally
Kinda Allaham	251518558	Contributed equally
Margie Pigott	251147800	Contributed equally

1) Rational for the tool/review

Canada faces prolonged dermatology wait times, averaging approximately five months (Canadian Dermatology Association, 2025). AI-based dermatology applications have emerged as scalable triage tools, with some demonstrating dermatologist-level accuracy (Esteva et al., 2017). However, these systems are frequently trained on image datasets that overrepresent lighter skin tones.

Canada faces prolonged dermatology wait times, averaging approximately five months (Canadian Dermatology Association, 2025). AI-based dermatology applications have emerged as scalable triage tools, with some demonstrating dermatologist-level accuracy (Esteva et al., 2017). However, these systems are frequently trained on image datasets that overrepresent lighter skin tones.

Empirical studies demonstrate significant performance disparities across skin tones. One study reported AI diagnostic accuracy of approximately 70% in lighter skin types but only 17% in darker skin conditions (Kamulegeya et al., 2023). Dataset evaluations confirm uneven representation across Fitzpatrick skin categories (Groh et al., 2021). Broader research shows that biased algorithmic training can systematically disadvantage racialized populations when deployed at scale (Obermeyer et al., 2019).

In melanoma specifically, individuals with darker skin tones are more likely to present at advanced stages, contributing to higher mortality (Thompson et al., 2023). Part of this disparity is linked to differences in lesion presentation that are not adequately represented in training images.

Current AI tools often treat melanoma diagnostic features as visually uniform across skin tones. However, the clinical appearance of ABCDE features (Asymmetry, Border irregularity, Color variation, Diameter, Evolving) can differ significantly depending on background pigmentation (Tsao et al., 2015; Thompson et al., 2023).

There is currently no structured diagnostic criteria model that stratifies melanoma indicators by skin tone and feeds those criteria into AI systems.

2) Purpose and aims for the tool/review

This project proposes a skin tone stratified diagnostic criteria model for integration into AI dermatology applications. Figure 1 displays a visual description of the various aims and goals.

The aims are to:

- Develop ABCDE-based melanoma criteria adapted for distinct skin tone ranges.

- Translate clinically observed presentation differences into structured diagnostic parameters.
- Integrate these stratified criteria into AI training datasets.
- Improve diagnostic sensitivity across pale, light, medium, and darker skin tones.

The ultimate goal is to operationalize equity in AI dermatology by ensuring diagnostic criteria reflect biological and phenotypic variation.

Image 1. Skin Tone Stratified Diagnostic Framework to Guide AI

3) Target audience/field of research for the tool/review

The primary users and experts of this model include AI developers and biomedical engineers designing dermatology diagnostic tools, as well as dermatologists and dermatopathologists who contribute to image labeling and clinical validation of datasets. Patients who use AI triage applications are also directly affected, as the accuracy of these tools can influence diagnostic recommendations and healthcare outcomes. A visual description can be found in Figure 2 a.

Secondary interest holders play an important part in maintaining safety and integrity to all parties and aspects working in the background to keep everything afloat. Secondary interest holders include research ethics boards, which oversee the ethical recruitment of participants, consent processes, and the responsible use of research data. Regulatory agencies also play an important role in evaluating the safety, accuracy, and potential bias of AI systems before they are implemented in clinical settings. In addition, policymakers in digital health governance develop regulations that guide the equitable and responsible use of AI in healthcare, helping ensure that emerging diagnostic technologies incorporate principles of equity, transparency, and accountability. A visual description can be found in Figure 2 b.

Figure 2. Target Audience (Retrieved from Google Images)

4) Final content and visuals for tool/review/guides

The skin tone stratified diagnostic criteria model consists of three integrated components:

A. Skin Tone Stratification Categories

The traditional Fitzpatrick Skin Type Scale, displayed in Figure 3, is widely used in dermatology, but it was originally designed to classify skin based on how it responds to ultraviolet radiation, particularly how easily it burns or tans. Because of this, it does not fully represent the diversity of global skin tones or how dermatological conditions can appear across different populations. To address this limitation, a more comprehensive and representative classification model could expand the existing framework while maintaining the familiar Fitzpatrick nomenclature.

In this updated model, shown in Figure 4, the Fitzpatrick categories could be supplemented with additional descriptors that better capture variations in pigmentation and disease presentation across skin tones. This would improve inclusivity and help support more accurate diagnoses in both clinical dermatology and AI-based diagnostic tools.

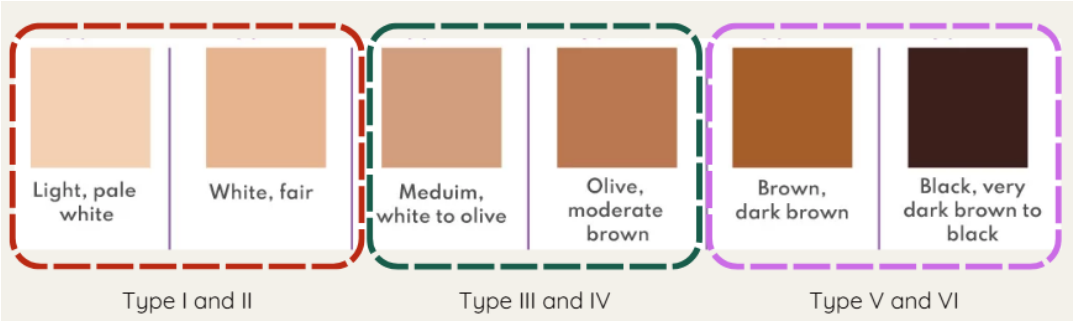


Figure 3. Fitzpatrick Scale (Retrieved from Google Images)

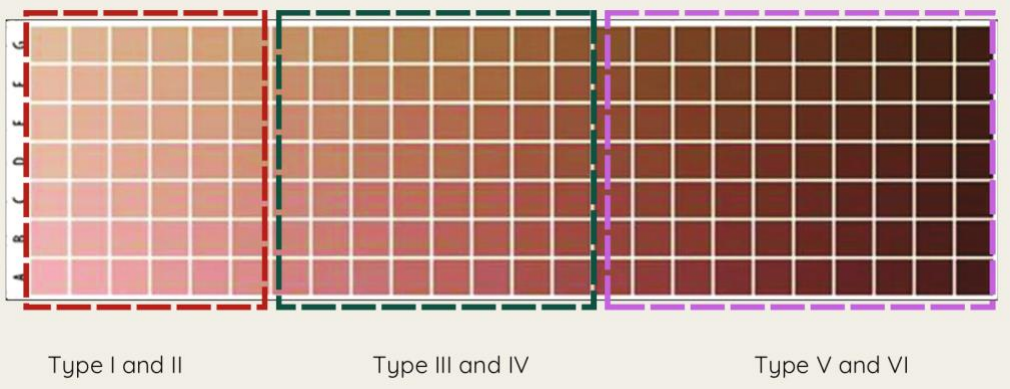


Figure 4. Comprehensive Skin Scale (Retrieved from Google Images)

B. ABCDE Criteria Adapted by Skin Tone

The model builds on the established ABCDE framework (Tsao et al., 2015) but modifies feature interpretation by skin tone.

Original Framework:

- **A – Asymmetry:** One half of the spot is unlike the other half.
- **B – Border:** The spot has an irregular, scalloped, or poorly defined border.
- **C – Color:** The spot has varying colors from one area to the next (tan, brown, black, white, red, or blue).
- **D – Diameter:** The spot is usually greater than 6 mm (about the size of a pencil eraser), but melanomas can be smaller.
- **E – Evolving:** The spot is changing in size, shape, color, or appearance over time.

Adapted Framework:

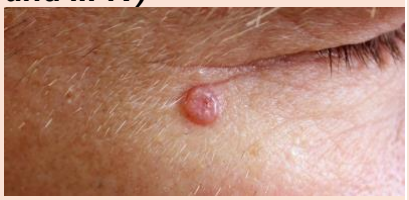
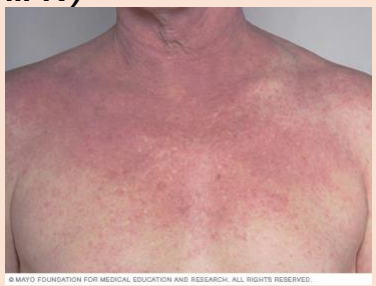
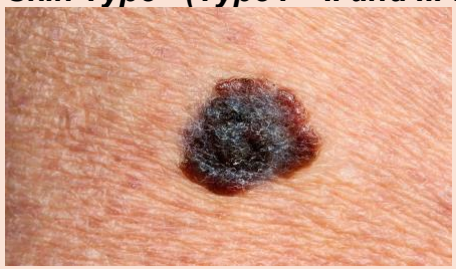
- A – Appearance
- B – Border
- C – Colour
- D – Damage
- E – Evolution

Our adapted ABCDE framework expands the traditional melanoma screening criteria so it can be applied to a wider range of dermatological conditions and diverse skin tones. While the original system focuses on features specific to melanoma progression, the modified version has adapted some elements to make the criteria more broadly applicable to other skin diseases that may not show the same morphological

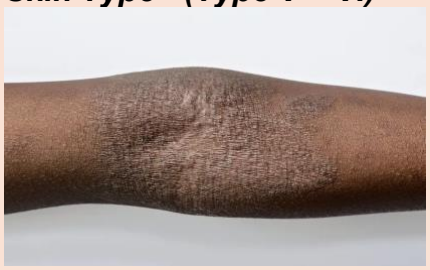
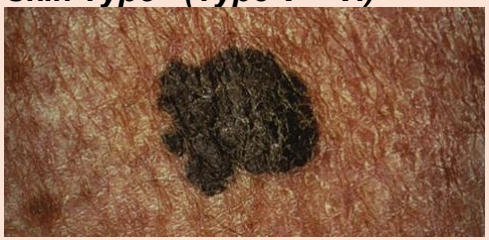
changes. Table 1 shows a more comprehensive description of the various diseases. A was adapted from Asymmetry to Appearance, allowing clinicians and AI tools to assess overall visual characteristics of a lesion, such as texture, elevation, or unusual presentation, rather than relying solely on symmetry. D was changed from Diameter to Damage, which focuses on signs of skin disruption such as ulceration, irritation, bleeding, or tissue breakdown that may occur in many dermatological conditions beyond melanoma. The remaining criteria, Border, Colour, and Evolution, continue to capture important visual changes in lesion structure, pigmentation, and progression over time. Together, this adapted framework allows for a more inclusive and flexible diagnostic tool that can support both clinicians and AI dermatology applications when evaluating a broader range of skin conditions.

Table 1. Curated Skin Tone Stratified Diagnostic Criteria

Disease	Basal Cell Carcinoma	Atopic Dermatitis (Eczema)	Melanoma
Description of Disease	The most common type of skin cancer that occurs mainly through UV radiation exposure leading to DNA damage (Skin Cancer Foundation, 2019).	A chronic condition characterized by dry, itchy, and inflamed skin. (Mayo Clinic, 2023).	Type of skin cancer that originates from the melanocytes. Mainly caused by exposure to UV lights (Mayo Clinic, 2023b)
Skin Type - (Type I – II and III-IV)			
A (Appearance)	• Pink or red patches • Pearly or translucent nodules	• Dry, rough, scaly skin • Oozing or vesicles in severe flares • Red or dark pink inflamed patches	• One half of the lesion does not match the other half • Irregular elevation (raised in one area, flat in another)
B (Border)	• Slightly raised, rolled, or pearly borders • Borders may appear shiny or waxy	• Poorly demarcated patches • Typical flexural distribution (elbows, knees, neck, wrists)	• Irregular, jagged, blurred, or frayed borders
C (Colour)	• Pink, red or flesh colour	• **Hyper-pigmentation:** • Darker patches during or after flares • **Hypo-pigmentation:** • Lighter patches as inflammation resolves	• Brown, black, blue, red, or white areas in one lesion • Often >5 mm, but not always
D (Damage)	• Central indentation or ulceration	• Lichenification (thickened,leathery skin with exaggerated skin lines)	• Ulceration or open sores on the lesion

	• Crusting or surface breakdown • May ooze, bleed, or scab repeatedly	• Dark pink or red thickened plaques	• Bleeding or oozing without injury • Crusting or scabbing that does not heal • Surface becoming raised, rough, or eroded • Pain or tenderness developing in a previously painless mole
E (Evolution)	• Lesion that itches, bleeds, crusts, or oozes • Recurrent non-healing sore • Gradual enlargement over time	• Chronic, relapsing course • Intense itch, often preceding visible changes • Pigmentary changes may persist months to years after inflammation resolve	• Change in size, shape, color, or surface • New symptoms: itching, bleeding, crusting • New lesion that looks different from others (“ugly duckling”) • Change over weeks to months
Image Guide	• ***Skin Type - (Type I – II and III-IV)***  *Finding Skin Cancer in Darker Skin Tones*, 2025	• ***Skin Type - (Type I – II and III-IV)***  *Atopic Dermatitis (Eczema) - Symptoms and Causes*, 2025	• ***Skin Type - (Type I – II and III-IV)***  Peri, 2022
Skin Type - (Type V – VI)			
A (Appearance)	• Pigmented lesions (tan, brown, or black)	• Inflammation appears as hyper-pigmented or dull brown/grey areas	• One half does not match the other • Irregular elevation (raised in one area, flat in another)

	• May resemble a benign mole → higher risk of under-recognition	• Follicular prominence (small itchy papules), especially on trunk and forearms • Dry scale may look white or grey, making flares more visible	• Asymmetry may be subtle if colour contrast is low → shape matters more
B (Border)	• Slightly raised, rolled, or pearly borders • Borders may appear shiny or waxy • Rolled edges may be visible without surrounding redness	• Poorly demarcated patches • Typical flexural distribution (elbows, knees, neck, wrists) • Borders may blend into surrounding pigmentation, masking severity	• Irregular, jagged, blurred, or frayed borders • Borders may blend into surrounding pigmentation • Nail melanoma may show uneven pigment spreading into cuticle (Hutchinson sign)
C (Colour)	• Frequently uniformly pigmented (brown, black, tan) • Can be mistaken for a stable nevus	• **Hyper-pigmentation:** o Darker patches during or after flares o More pronounced and persistent in darker skin • **Hypopigmentation:** o Lighter patches as inflammation resolves o Seen in all skin types, but more noticeable in darker skin	• May appear only *slightly darker* than surrounding skin • Can be pink, red, colorless, or uniformly dark • Can mimic benign pigmentation (e.g., nail streaks, lip patches)
D (Damage)	• Central indentation or ulceration • Crusting or surface breakdown	• **Lichenification** (thickened,leathery skin with exaggerated skin lines) • Slate grey or dark brown thickened areas	• Ulceration or bleeding in areas not usually monitored (palms, soles, genitals)

	• May ooze, bleed, or scab repeatedly	• Common on wrists, forearms, neck, lower legs, genitals	• Nail involvement: nail splitting, lifting, destruction, or bleeding under the nail • Persistent cracking or breakdown mistaken for trauma or infection • Pain, tenderness, or bleeding without clear cause • Subtle surface erosion without obvious color change
E (Evolution)	• Lesion that itches, bleeds, crusts, or oozes • Recurrent non-healing sore • Gradual enlargement over time	• Chronic, relapsing course • Intense itch, often preceding visible changes • Pigmentary changes may persist months to years after inflammation resolve • Pigmentary sequelae may be more distressing than active eczema	• Change in size, shape, color, or surface • New symptoms: itching, bleeding, crusting • New lesion that looks different from others (“ugly duckling”) • Change over weeks to months
Image Guide	***Skin Type - (Type V – VI)***  *Finding Skin Cancer in Darker Skin Tones*, 2025	***Skin Type - (Type V – VI)***  Frysh, 2022	***Skin Type - (Type V – VI)***  *Pictures of Melanoma: Skin Changes to Look for and More*, 2014
Citation	Skin Cancer Foundation, 2019	Lambert, 2021	*Melanoma: Learn More – Detecting Melanoma*, 2024

C. AI Integration Mechanism

The adapted criteria are:

1. Used to guide image labeling during dataset creation.
2. Structured as metadata inputs linked to each image.
3. Incorporated into supervised machine learning training.

This allows the AI system to learn:

- How ABCDE features manifest differently across skin tones.
- When diagnostic confidence should be lowered.
- When referral flags should be triggered.

Instead of assuming uniform feature presentation, the AI is trained to recognize stratified diagnostic variability.

5) Implications and implementation of the tool/review

The proposed skin tone stratified diagnostic framework has important implications for the development and deployment of AI dermatology tools. Current machine learning models in dermatology have demonstrated uneven performance across skin tones due to imbalanced training datasets that disproportionately represent lighter skin types (Groh et al., 2021; Kamulegeya et al., 2023). By incorporating clinically validated diagnostic criteria that account for variation in lesion presentation across skin tones, AI systems can be trained to detect dermatologic conditions more accurately in diverse populations. This approach addresses known diagnostic disparities and may improve model sensitivity while reducing false negative rates in populations that have historically been underrepresented in dermatologic datasets (Adamson & Smith, 2018).

Implementation of the model would involve collaboration between AI developers, dermatologists, and dermatopathologists. Dermatology clinics and pathology laboratories would contribute clinically validated images representing a broad spectrum of skin tones and disease presentations. These images would then be labeled using the adapted ABCDE diagnostic framework, which includes appearance, border irregularity, color variation, damage, and evolution (Tsao et al., 2015). By integrating these diagnostic criteria into dataset labeling and metadata, machine learning systems can be trained to recognize how dermatologic conditions manifest differently across skin pigmentation levels.

During model training, these criteria would function as structured inputs that guide supervised learning algorithms in identifying diagnostic features across stratified skin tone groups. Continuous monitoring and evaluation of model performance would

also be necessary to ensure equitable accuracy across populations. Algorithms that demonstrate reduced performance in specific demographic groups could be flagged for retraining using additional representative data. This repeating process helps reduce algorithmic bias and supports the development of more reliable and inclusive diagnostic technologies (Obermeyer et al., 2019). Through these steps, the framework can be integrated into existing dermatology AI systems as a structured mechanism for improving fairness, accuracy, and clinical validity.

6) Wider impact of tool/review

The broader impact of this framework extends beyond improving the technical performance of AI dermatology applications. By incorporating skin tone specific diagnostic criteria into AI training systems, the model helps address longstanding inequities in dermatologic diagnosis and medical imaging datasets, this is shown in Figure 5. Historically, dermatologic education and datasets have overrepresented lighter skin tones, contributing to diagnostic challenges and delayed detection in patients with darker skin tones (Adamson & Smith, 2018). Integrating diverse diagnostic criteria into AI systems therefore supports improved diagnostic equity and helps ensure that emerging healthcare technologies are developed with greater representational balance and clinical inclusivity.

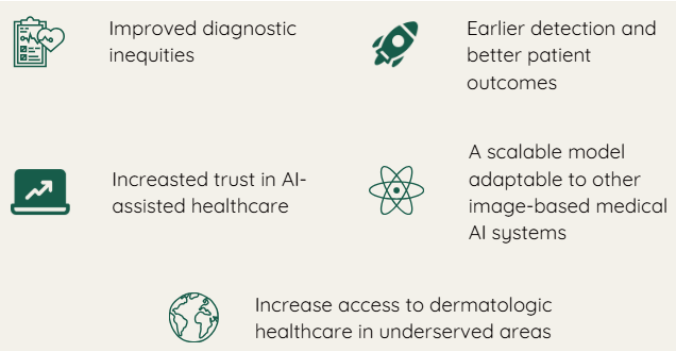


Figure 5. Wider Impact of the skin tone stratified AI dermatology framework

Improving diagnostic accuracy across diverse populations may also support earlier detection of serious dermatologic conditions such as melanoma. Individuals with darker skin tones are more likely to be diagnosed at later stages of melanoma, which is associated with poorer clinical outcomes and higher mortality rates (Thompson et al., 2023). By enabling AI-assisted screening tools to recognize suspicious lesions across a wider range of skin tones, the proposed framework may contribute to earlier clinical evaluation, improved treatment outcomes, and better overall patient care.

In addition to improving clinical outcomes, the framework supports more responsible development of medical artificial intelligence. Embedding equity considerations into dataset development, labeling processes, and algorithm training helps ensure that bias mitigation becomes a standard component of AI design (Obermeyer et al., 2019). As AI systems become more equitable and transparent, this may also increase trust in AI-assisted healthcare among both clinicians and patients.

Another important impact of this framework is its scalability. Because the model structures diagnostic criteria in a way that can be incorporated into machine learning datasets, it can be adapted for other image-based medical AI systems beyond dermatology. This approach demonstrates how diversity-aware diagnostic criteria can be systematically integrated into algorithm development across medical imaging fields.

Finally, improved reliability of AI-assisted dermatology tools may help increase access to dermatologic care in underserved or resource-limited settings. In Canada, dermatology wait times can exceed several months, limiting timely access to specialist care (Canadian Dermatology Association, 2025). AI-supported triage tools with improved diagnostic accuracy could help prioritize urgent cases, reduce unnecessary specialist consultations, and expand access to dermatologic screening in regions with limited specialist availability.

References

Atopic dermatitis (eczema) - Symptoms and causes. (2025). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/atopic-dermatitis-eczema/symptoms-causes/syc-20353273#dialogId42003413

Adamson, A. S., & Smith, A. (2018). Machine learning and health care disparities in dermatology. *JAMA Dermatology, 154*(11), 1247–1248. https://doi.org/10.1001/jamadermatol.2018.2348

Canadian Dermatology Association. (2025). *More than skin deep: 5 solutions to managing Canada's dermatologist shortage and preventing wider patient impacts*. https://dermatology.ca/wp-content/uploads/2025/01/CDA-Advocacy-Platform-EN-2.pdf

Finding skin cancer in darker skin tones. (2025). Aad.org. https://www.aad.org/public/diseases/skin-cancer/find/skin-of-color

Frysh, P. (2022, March 16). *How Eczema Affects Darker Skin Tones*. WebMD. https://www.webmd.com/skin-problems-and-treatments/eczema/eczema-darker-skin-tones

Groh, M., Harris, C., Soenksen, L., Lau, F., Han, R., Kim, A., Koochek, A., Badri, O., & Rajpurkar, P. (2021). Evaluating deep neural networks trained on clinical images in dermatology with the Fitzpatrick 17k dataset. *Nature Medicine, 27*(9), 165–171. https://doi.org/10.48550/arXiv.2104.09957

Kamulegeya, L., Bwanika, J., Okello, M., Rusoke, D., Nassiwa, F., Lubega, W., Musinguzi, D., & Börve, A. (2023). Using artificial intelligence on dermatology conditions in Uganda: A case for diversity in training data sets for machine learning. *African Health Sciences, 23*(2), 753–763. https://doi.org/10.4314/ahs.v23i2.86

Lambert, A. (2021, February 9). *Skin pigmentation and eczema*. National Eczema Society. https://eczema.org/information-and-advice/living-with-eczema/skin-pigmentation/

Mayo Clinic. (2023). *Atopic dermatitis (eczema) - Symptoms and causes*. Atopic Dermatitis (Eczema); Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/atopic-dermatitis-eczema/symptoms-causes/syc-20353273

Mayo Clinic. (2023b, December 30). *Melanoma - Symptoms and Causes*. Mayo Clinic; Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/melanoma/symptoms-causes/syc-20374884

Melanoma: Learn More – Detecting melanoma. (2024, February 21). Nih.gov; Institute for Quality and Efficiency in Health Care (IQWiG). https://www.ncbi.nlm.nih.gov/books/NBK321126/

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Peri, C. (2022, April 26). *Metastatic Melanoma*. WebMD. https://www.webmd.com/melanoma-skin-cancer/metastatic-melanoma

Pictures of Melanoma: Skin Changes to Look for and More. (2014, March 25). Healthline. https://www.healthline.com/health/melanoma-pictures

Skin Cancer Foundation. (2019). *Basal Cell Carcinoma*. The Skin Cancer Foundation. https://www.skincancer.org/skin-cancer-information/basal-cell-carcinoma/

Thompson, B., Jenkins, T., Sanchez, J., Frederick, M., Posligua-Alban, A., & Barbosa, N. S. (2023). Melanoma: Does it present differently in darker skin tones? *MedEdPORTAL*. https://doi.org/10.15766/mep_2374-8265.11311

Tsao, H., Olazagasti, J. M., Cordoro, K. M., Brewer, J. D., Taylor, S. C., Bordeaux, J. S., Chren, M.-M., Sober, A. J., Tegeler, C., Bhushan, R., & Begolka, W. S. (2015).

Early detection of melanoma: Reviewing the ABCDEs. *Journal of the American Academy of Dermatology, 72*(4), 717–723. https://doi.org/10.1016/j.jaad.2015.01.025