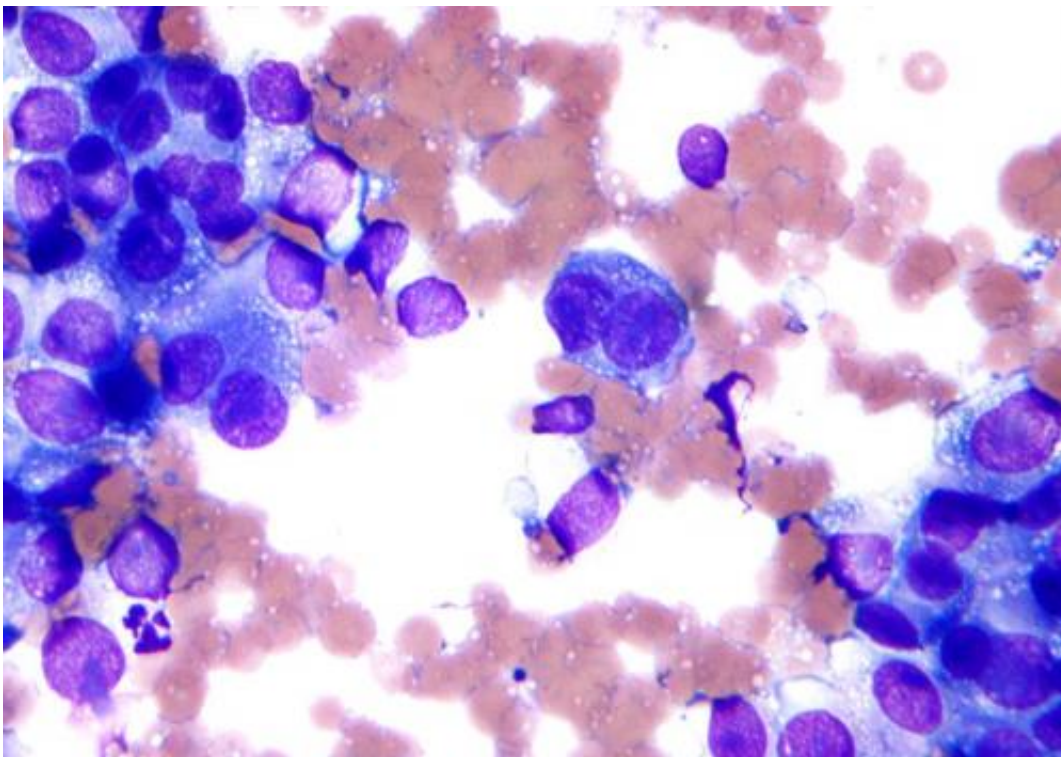


Data available for training AI to spot skin cancer are insufficient and lacking in pictures of darker skin

November 10 2021



Micrograph of malignant melanoma. Cytology specimen. Field stain. Credit: Nephron/Wikipedia

The images and accompanying data available for training artificial intelligence (AI) to spot skin cancer are insufficient and include very few images of darker skin, according to research presented at the NCRI

Festival and published in *Lancet Digital Health*.

AI is increasingly being used in medicine as it can make diagnosis of diseases like [skin](#) cancer quicker and more effective. However, AI needs to be 'trained' by looking at data and images from a large number of patients where the diagnosis has already been established and so an AI program depends heavily upon the information it is trained on.

Researchers say there is an urgent need for better sets of data on [skin cancers](#) and other skin lesions which contain information on who is represented in the datasets.

The research was presented by Dr. David Wen from the University of Oxford, UK. He said: "AI programs hold a lot of potential for diagnosing skin cancer because it can look at pictures and quickly and cost-effectively evaluate any worrying spots on the skin. However, it's important to know about the images and patients used to develop programs, as these influence which groups of people the programs will be most effective for in real-life settings. Research has shown that programs trained on images taken from people with lighter skin types only might not be as accurate for people with darker skin, and vice versa."

Dr. Wen and his colleagues carried out the first ever review of all freely accessible sets of data on [skin lesions](#) around the world. They found 21 sets including more than 100,000 pictures.

Diagnosis of skin cancer normally requires a photo of the worrying lesion as well as a [picture](#) taken with a special hand-held magnifier, called a dermatoscope, but only two of the 21 datasets included images taken with both of these methods. The datasets were also missing other important information, such as how images were chosen to be included, and evidence of ethical approval or patient consent.

Fourteen of 21 datasets gave information on which country they came from and of those, nine contained images from European countries. Only a small percentage of images were accompanied by information about the patients' skin colour or ethnicity. Among pictures where skin colour was stated (2,436 pictures), only ten were of brown skin and only one was of dark brown or black skin. Among pictures where ethnicity was stated (1,585 pictures), none were from people with African, Afro-Caribbean or South Asian background.

Dr. Wen added: "We found that for the majority of datasets, lots of important information about the images and patients in these datasets wasn't reported. There was limited information on who, how and why the images were taken. This has implications for the programs developed from these [images](#), due to uncertainty around how they may perform in different groups of people, especially in those who aren't well represented in datasets, such as those with darker skin. This can potentially lead to the exclusion or even harm of these groups from AI technologies.

"Although skin cancer is rarer in people with darker skins, there is evidence that those who do develop it may have worse disease or be more likely to die of the disease. One factor contributing to this could be the result of skin cancer being diagnosed too late."

To guard against this, Dr. Wen and his colleagues hope to create quality standards for health data used in AI development. This will include information on who should be represented in datasets and which patient characteristics should be recorded.

Dr. Neil Steven is a member of the NCRI Skin Group, Honorary Consultant in Medical Oncology at University Hospitals Birmingham NHS Foundation Trust, UK and was not involved in the research. He said: "Skin cancer affects more than 200,000 people each year in the UK

alone. Some types of skin cancer are more aggressive than others so quick diagnosis and treatment can be vital.

"We already know that there are not enough pictures of people from black and Asian backgrounds in the textbooks we use to train doctors. The findings of this review—that pictures of people with darker skin are under-represented in datasets—raise concerns about the ability of AI to assist in skin [cancer](#) diagnosis, especially in a global context.

"I hope this work will continue and help ensure that the progress we make in using AI in medicine will benefit all patients, recognising that human [skin colour](#) is highly diverse."

More information: David Wen et al, Characteristics of publicly available skin cancer image datasets: a systematic review, *The Lancet Digital Health* (2021). [DOI: 10.1016/S2589-7500\(21\)00252-1](https://doi.org/10.1016/S2589-7500(21)00252-1)

Provided by National Cancer Research Institute

Citation: Data available for training AI to spot skin cancer are insufficient and lacking in pictures of darker skin (2021, November 10) retrieved 21 June 2024 from <https://medicalxpress.com/news/2021-11-ai-skin-cancer-insufficient-lacking.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.