Algorithm-based smartphone apps to assess risk of skin cancer in adults: critical appraisal of a systematic review*

DOI: 10.1111/bjd.19502

Up to three million skin cancers occur worldwide annually, with an increasing burden over time.¹ Melanomas, accounting for 4% of skin cancers, are responsible for 75% of deaths.¹,² An estimated 86% of melanomas in the UK are preventable.³ Self-monitoring, earlier detection and improvement in survival from melanoma are therefore of interest. Algorithm-based smartphone applications (apps) are based on artificial intelligence algorithms trained on images of diagnosed skin cancer. They have the potential to facilitate early detection of otherwise aggressive life-threatening skin cancers, thus prompting users to seek timely medical attention and improve survival.

Freeman et al. conducted a systematic review investigating the accuracy of smartphone apps used for screening for skin cancer in the general population.⁴ Nine studies in total met the inclusion criteria. Six studies verified their results using histology or follow-up (n = 725), with five studies evaluating the detection of melanoma and one study evaluating the SkinVision app, aiming to detect other skin cancers such as basal cell carcinoma and squamous cell carcinoma as well as premalignant conditions. Three studies, all evaluating the SkinVision app, verified their results by expert recommendations for whether further investigation or intervention should be pursued (n = 407), with two including diagnoses of all types of premalignant and malignant lesions and one study not reporting the expert diagnosis (Ngoo et al. 2018). The results of the paper overall show variable and unreliable test accuracy for the six apps (Table 1).

This systematic review was well conducted and reported. The authors used QUADAS-2 to assess study quality, which is a validated tool to assess diagnostic accuracy studies.⁵ The restriction of studies to those using smartphone assessment ensures that the results will be generalizable to a setting based on population screening. The authors were not funded by sponsors with a financial interest in skin-cancer-screening smartphone applications, meaning bias favourable to industry sponsors is unlikely. Neither the protocol nor the manuscript reported on whether the authors reviewed the sources of funding for the studies included in the review, as recommended in AMSTAR-2.⁶ Although this item does not form part of the QUADAS-2 tool, evaluation studies funded or performed by the industry sponsor may introduce bias in favour of the smartphone app. We have summarized the sources of funding for the included studies in Table 1.

Each included study had a high risk of bias in at least one of four QUADAS-2 domains. The studies included in the review selectively recruited participants who were not fully representative of the general population. Seven studies recruited patients already selected for excision of suspicious lesions, or

| Study          | App(s)                                      | Sources of funding                      | Conflicts of interest declared                                      |
|----------------|---------------------------------------------|-----------------------------------------|--------------------------------------------------------------------|
| Robson 2012    | MelApp                                      | None                                    | None declared                                                      |
| Wolf 2013      | Apps (three) not named                      | National Institutes of Health grant     | Yes: author was investigator and consultant for MELA Sciences, Inc.|
| Chadwick 2014  | SkinScan, MelApp, Mole Detective, Spot Mole Plus, Dr Mole Premium | Not stated                             | Not stated                                                         |
| Maier 2015     | SkinVision                                  | SkinVision                             | Yes: two authors were consultants for SkinVision                   |
| Dorairaj 2017  | App (one) not named                         | Not stated                             | None declared                                                      |
| Nabil 2017     | SkinVision                                  | Not stated                             | Not stated                                                         |
| Thissen 2017   | SkinVision                                  | SkinVision                             | Yes: two authors received fees                                     |
| Chung 2018     | SkinVision                                  | None                                   | None declared                                                      |
| Ngoo 2018      | SkinVision, SpotMole, Dr Mole               | National Health and Medical Research Council grant | Yes: author is a shareholder of e-derm consult GmbH and MoleMap by Dermatologists Ltd Pty |

Table 1 The sources of funding and conflicts of interest of each study included in the original review, and the related smartphone app(s) where relevant⁴
undergoing assessment of their skin conditions by dermatologists.\textsuperscript{4} It was unclear whether skin of colour and special sites, such as acral skin and nails, were represented adequately. Without a focus on whether the patient sample is generalizable to the population, including an adequate representation of different skin colours and body sites, algorithms risk producing biased results. Given that acral melanoma disproportionately affects people of darker skin, poor representation of the above in the training dataset may potentially worsen healthcare disparities.\textsuperscript{7} Other major limitations highlighted by the authors include inadequate reference standards, differential verification and high rates of unreliable images. It should be noted that the test characteristics of SkinVision for high-risk lesions are not specific to any one type of skin cancer; it is therefore unclear what the sensitivity and specificity of the algorithm are in relation to any particular type of skin cancer.

The positive predictive value (PPV) of the algorithms was not an item reported or extracted in this systematic review. This characteristic of a test, when evaluated in a representative sample for the target population, is important for an indication of the usefulness of the test as a screening tool. Any given screening test will likely have a higher PPV for common skin cancers, and have a lower PPV with a high false positive rate in a disease of low prevalence, such as melanoma. This may have a low impact on improving patient survival from melanoma but a high adverse effect of increasing overdiagnosis, unnecessary healthcare visits and potentially procedures, increasing the use of scarce healthcare resources.

Freeman et al. produced a high-quality systematic review, showing that there is little evidence that smartphone apps can be useful in self-monitoring and detection of skin cancer at present. We suggest that future studies and systematic reviews involving smartphone apps for skin cancer diagnosis should report the presence of industry funding, report the PPV where appropriate, highlight whether images of special body sites were included in the training and validation datasets, and report whether the proportion of participants with skin of colour reflects the target population.

I.S. Malhi\textsuperscript{1} and Z.Z.N. Yiu\textsuperscript{2}
\textsuperscript{1}Manchester Medical School, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK; and \textsuperscript{2}Dermatology Centre, Salford Royal NHS Foundation Trust, NIHR Manchester Biomedical Research Centre, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
Correspondence: Zenas Z.N. Yiu.
Email: zenas.yiu@manchester.ac.uk

*Plain language summary available online

References

1 World Health Organization. Ultraviolet (UV) radiation and skin cancer. Available at: https://www.who.int/news-room/q-a-detail/ultraviolet-(uv)-radiation-and-skin-cancer (last accessed 14 September 2020).
2 Laikova KV, Oberemok VV, Krasnodubets AM et al. Advances in the understanding of skin cancer: ultraviolet radiation, mutations, and antisense oligonucleotides as anticancer drugs. Molecules 2019; \textit{24}:1516.
3 Cancer Research UK. Melanoma skin cancer statistics. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer (last accessed 14 September 2020).
4 Freeman K, Dinnes J, Chuchu N et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ 2020; \textit{368}:m127.
5 Whiting P. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; \textit{155}:529.
6 Shea B, Reeves B, Wells G et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017; \textit{358}:j4008.
7 Adamson A, Smith A. Machine learning and health care disparities in dermatology. JAMA Dermatol 2018; \textit{154}:1247.

Funding sources: none.

Conflicts of interest: The authors declare they have no conflicts of interest.