

Abstract

Testing Artificial Intelligence Algorithms in the Real World: Lessons From the SMARTI Trial

Claire Felmingham^{1,2}; Gabrielle Byars³; Simon Cumming³; Jane Brack²; Zongyuan Ge⁴; Samantha MacNamara⁵; Martin Haskett⁵; Rory Wolfe¹; Victoria Mar^{1,2}

¹School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

²Victorian Melanoma Service, Alfred Hospital, Melbourne, Australia

³Melanoma and Skin Cancer Research Centre, Monash University, Melbourne, Australia

⁴Monash eResearch Centre, Monash University, Melbourne, Australia

⁵Molemap Ltd, Melbourne, Australia

Corresponding Author:

Victoria Mar

School of Public Health and Preventive Medicine

Monash University

553 St Kilda Road

Melbourne, VIC 3004

Australia

Phone: 61 3 9903 0556

Email: Victoria.Mar@monash.edu

Abstract

Background: A number of studies have shown promising performance of artificial intelligence (AI) algorithms for diagnosis of lesions in skin cancer. To date, none of these have assessed algorithm performance in the real-world setting.

Objective: The aim of this project is to evaluate practical issues of implementing a convolutional neural network developed by MoleMap Ltd and Monash University eResearch in the clinical setting.

Methods: Participants were recruited from the Alfred Hospital and Skin Health Institute, Melbourne, Australia, from November 1, 2019, to May 30, 2021. Any skin lesions of concern and at least two additional lesions were imaged using a proprietary dermoscopic camera. Images were uploaded directly to the study database by the research nurse via a custom interface installed on a clinic laptop. Doctors recorded their diagnosis and management plan for each lesion in real time. A pre-post study design was used. In the preintervention period, participating doctors were blinded to AI lesion assessment. An interim safety analysis for AI accuracy was then performed. In the postintervention period, the AI algorithm classified lesions as benign, malignant, or uncertain after the doctors' initial assessment had been made. Doctors then had the opportunity to record an updated diagnosis and management plan. After discussing the AI diagnosis with the patient, a final management plan was agreed upon.

Results: Participants at both sites were high risk (for example, having a history of melanoma or being transplant recipients). 743 lesions were imaged in 214 participants. In total, 28 dermatology trainees and 17 consultant dermatologists provided diagnoses and management decisions, and 3 experienced teledermatologists provided remote assessments. A dedicated research nurse was essential to oversee study processes, maintain study documents, and assist with clinical workflow. In cases where AI algorithm and consultant dermatologist diagnoses were discordant, participant anxiety was an important factor in the final agreed management plan to biopsy or not.

Conclusions: Although AI algorithms are likely to be of most use in the primary care setting, higher event rates in specialist settings are important for the initial assessment of algorithm safety and accuracy. This study highlighted the importance of considering workflow issues and doctor-patient-AI interactions prior to larger-scale trials in community-based practices.

Acknowledgments: This research was supported by the Victorian Medical Research Acceleration Fund, with 1:1 contribution from MoleMap Ltd. VM is supported by the National Health and Medical Research Council Early Career Fellowship. CF is supported by the Monash University Research Training Program Scholarship.

Conflicts of Interest: SM is head of clinical research and regulatory affairs at Kahu.ai Ltd, a subsidiary of MoleMap Ltd. MH was the chief medical officer and a director of MoleMap Ltd, and holds shares in MoleMap Ltd.

Trial Registration: ClinicalTrials.gov NCT04040114; <https://clinicaltrials.gov/ct2/show/NCT04040114>

(*iproc* 2022;8(1):e36902) doi: [10.2196/36902](https://doi.org/10.2196/36902)

KEYWORDS

melanoma; artificial intelligence; algorithm; dermatology; skin cancer

Multimedia Appendix 1

Study procedure.

[\[PNG File , 1104 KB-Multimedia Appendix 1\]](#)

Edited by T Derrick; this is a non-peer-reviewed article. Submitted 28.01.22; accepted 28.01.22; published 01.03.22.

Please cite as:

Felmingham C, Byars G, Cumming S, Brack J, Ge Z, MacNamara S, Haskett M, Wolfe R, Mar V

Testing Artificial Intelligence Algorithms in the Real World: Lessons From the SMARTI Trial

iproc 2022;8(1):e36902

URL: <https://www.iproc.org/2022/1/e36902>

doi: [10.2196/36902](https://doi.org/10.2196/36902)

PMID:

©Claire Felmingham, Gabrielle Byars, Simon Cumming, Jane Brack, Zongyuan Ge, Samantha MacNamara, Martin Haskett, Rory Wolfe, Victoria Mar. Originally published in *Iproceedings* (<https://www.iproc.org>), 01.03.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *Iproceedings*, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.iproc.org/>, as well as this copyright and license information must be included.