The impact of an artificial intelligence (AI) based app for skin cancer detection: a first clinical practice evaluation in a population-based setting

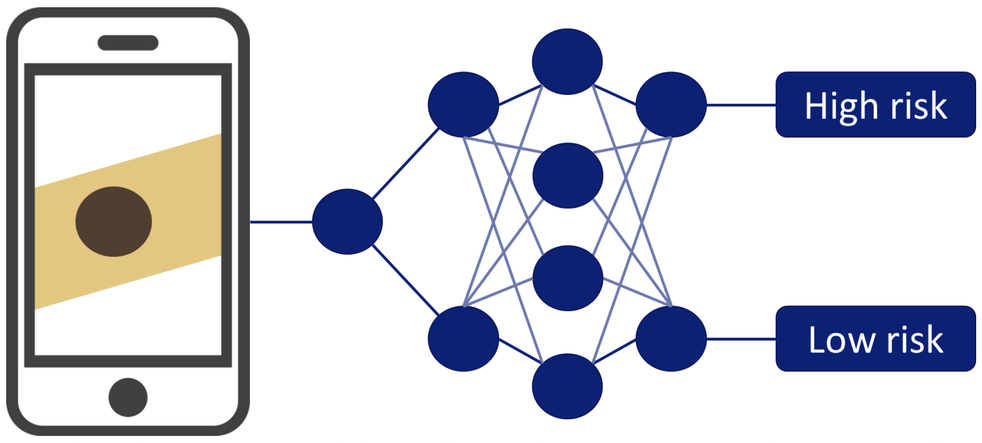
AM Smak Gregoor[1], TE Sangers[1], LB Bakker[2], LM Hollestein[1], CA Uyl-de Groot[2], TEC Nijsten[1], M Wakkee[1]

1 Department of Dermatology, Erasmus MC Cancer Institute Rotterdam, The Netherlands.

2 Erasmus School of Health Policy & Management, Erasmus University Rotterdam, The Netherlands.

Introduction & Objective

AI-based algorithms that can recognize skin cancer based on a photo have been implemented in mobile phone applications (mHealth).



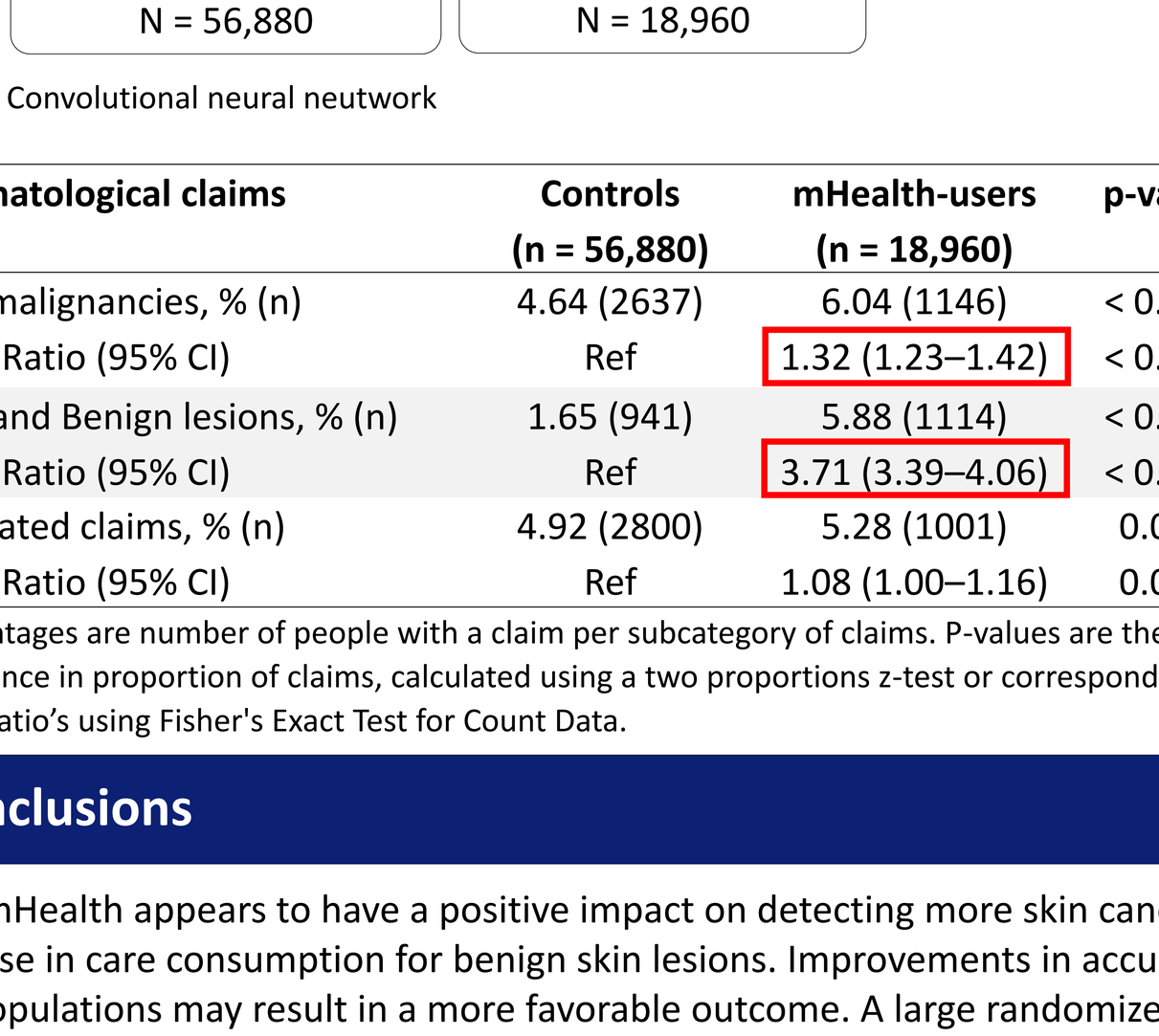
This study aims to provide a first glance on the impact of such an mHealth application for skin cancer detection on dermatological healthcare consumption in the Netherlands.

Methods

- **A retrospective population-based** pragmatic study.
- mHealth-users were matched 1:3 to non-mHealth-users.
- Differences in dermatological healthcare consumption were calculated based on **claims data** available from the **year 2019**.
- An early cost-effectiveness analysis was simulated for a range of sensitivities and specificities of the mHealth app. The incremental cost effectiveness ratio (ICER), was calculated using the following formula:

$$\text{ICER} = \frac{\text{Cost mHealth} - \text{Cost standard of care}}{\text{effectiveness mHealth} - \text{effectiveness standard of care}}$$

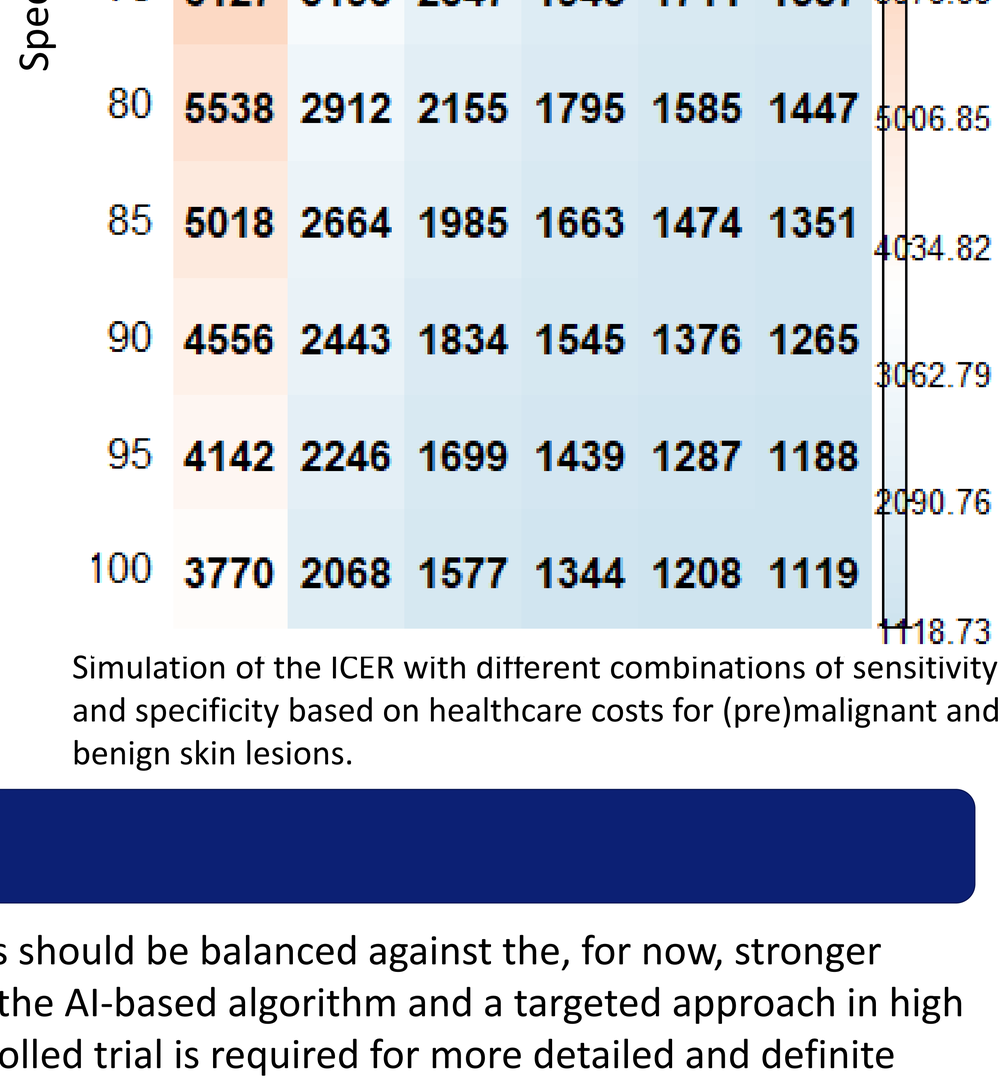
Results



*CNN; Convolutional neural neutwork

Dermatological claims	Controls (n = 56,880)	mHealth-users (n = 18,960)	p-value
(Pre)malignancies, % (n)	4.64 (2637)	6.04 (1146)	< 0.001
Odds Ratio (95% CI)	Ref	1.32 (1.23–1.42)	< 0.001
Nevi and Benign lesions, % (n)	1.65 (941)	5.88 (1114)	< 0.001
Odds Ratio (95% CI)	Ref	3.71 (3.39–4.06)	< 0.001
Unrelated claims, % (n)	4.92 (2800)	5.28 (1001)	0.066
Odds Ratio (95% CI)	Ref	1.08 (1.00–1.16)	0.066

Percentages are number of people with a claim per subcategory of claims. P-values are the difference in proportion of claims, calculated using a two proportions z-test or corresponding odds ratio's using Fisher's Exact Test for Count Data.



Simulation of the ICER with different combinations of sensitivity and specificity based on healthcare costs for (pre)malignant and benign skin lesions.

Conclusions

AI in mHealth appears to have a positive impact on detecting more skin cancer. This should be balanced against the, for now, stronger increase in care consumption for benign skin lesions. Improvements in accuracy of the AI-based algorithm and a targeted approach in high risk populations may result in a more favorable outcome. A large randomized controlled trial is required for more detailed and definite conclusions on its actual impact.

For questions, contact: ***m.wakkee@erasmusmc.nl***