PRINCIPLES AND PRACTICE OF CLINICAL RESEARCH
A Global Journal of Clinical Research

The Role of Smartphone-Based Applications and Artificial Intelligence Aiding in Screening, Faster Diagnosis and Prevention in Skin Cancer: A Systematic Mini-Review

Helena Brenes[1], Fernanda Maria de Queiroz Silva[1], Aisha Aladab[1], Abna Vieira[2], Hulya Ordek[1], Julia Minetto[3], Salim Al-Busaid[1], Adriana Margarita Rey[4], Catherine Catalan[1], Juan López-Cabello[5], Krisna Morales[1], Eolo Morandi[6], Karla Diaz[1], Layara Lipari[1], Taina Olivieri Chaves[1], Dolly Rojas[1], Alexia Perez[1], Ashraf Ahmed[1], Eduardo de Mello Araujo[1], Edgardo de Jesus Mateo[7], Merycarla Pichardo[8], Luiz Guilherme Darrigo Junior[9], Alvaro Guerra[1], Tim Jonas Hallenberger[10*]

[1] Harvard T.H Chan School of Public Health - ECPE, PPCR, Boston, USA; [2] Instituto do Câncer do Estado de São Paulo, São Paulo, Brazil; [3] Pediatric Liver Transplantation, Garrahan Hospital, Buenos Aires, Argentina; [4] Gastroenterology, University of Los Andes, Fundacion Santa Fe de Bogota, Bogota, Colombia; [5] Speech & Language Pathology Unit, Hospital del Trabajador, Santiago de Chile; [6] Instituto Butantan, São Paulo, São Paulo, Brazil; [7] Department of Neurosurgery, National Institute on Neurology and Neurosurgery MVS, Mexico City, Mexico; [8] URUS Urology, Santo Domingo, Dominican Republic; [9] Pediatric Bone Marrow Transplantation, Hapvida Medical Centre, Ribeirao Preto, Brazil; [10] Department of Neurosurgery, University Hospital Basel, Basel, Switzerland;

Abstract

Background: *Skin cancer, especially melanoma, is one of the most common cancers in humans. According to the WHO, it causes a significant burden on modern health services. Currently, diagnostics are painful and invasive for patients and healthcare providers, heralding novel techniques. In the age of mobile health, computer-aided applications could be an ideal solution. However, their relevance in clinical practice remains to be determined. This systematic review assessed detection rates and clinical implementation of computer-aided applications in patients with suspicious skin lesions.*

Methods and Results: *We systematically searched the following databases: PubMed, Medline, Embase (OVID), Web of Science, and the Cochrane Library for articles. Of 819 identified articles, eight remained for primary outcome analysis. Diagnostic sensitivity was reported in three out of eight studies, while accuracy and concordance were reported in three and four, respectively. The quality of the included studies was primarily moderate.*

Conclusion: *Although promising, computer-aided applications remain below standard detection rates of the current standard of care. Further improvement in accuracy and detection rates should be achieved while primary care providers and patients alike should be aware of the current limitations of these new diagnostic tools.*

Introduction

Skin cancer, most prominently melanoma, is one of the most common types of cancer, affecting around 325,000 patients in 2020 worldwide, thus posing a significant burden for healthcare systems (WHO, 2022).

Diagnostics for skin lesions classically consist of a visual clinical examination, a so-called face-to-face examination (FTF), with the help of a dermatoscopy. A specimen of the suspicious lesion must be collected via biopsy for a definitive diagnosis. As this is a rather lengthy and, for patients, often painful process, technological advances in medicine are sought to aid as additional diagnostic tools.

One of those current advances in technology-based platforms is smartphone applications and mobile Health (mHealth), mainly focusing on diagnostics and prevention in dermatology (WHO, 2022). However, their clinical relevance regarding implications

*Corresponding author: tim.hallenberger@gmail.com
Received: September 3, 2023 **Accepted:** November 6, 2023
Published: November 28, 2023
Editor: Felipe Fregni **Reviewers:** Pedro Faria, Dina Shehata, Niraj Mehta, Wendy Oliver, Ronny Enk, Rene Tovar-Parada
Keywords: melanoma, skin cancer, smartphone, artificial intelligence (AI), application
DOI: http://dx.doi.org/10.21801/ppcrj.2023.93.4

in clinical practice needs to be clarified.

Currently, only one extensive review of AI-based mHealth applications exists. Still, further information has yet to emerge to inform healthcare providers and patients on the ideal use of those applications, which, given the rapid evolution thereof, is essential (Freeman et al., 2020). Therefore, the purpose of the present work of conducting this study was to systematically review the current knowledge available regarding the use of technology-based applications, especially artificial intelligence and smartphone applications, to increase the detection of melanoma in high-risk populations.

Materials and Methods

This review was per the Preferred Reporting Items in Systematic Reviews and Meta-Analysis (PRISMA) guidelines. Search strings were translated using SR-Accelerator (Clark et al., 2020). The quality of the included studies was assessed using the Risk of Bias Tool 2 (RoB2) and the Risk of Bias in Non-Randomized Studies (ROBINS) tool (J. A. Sterne et al., 2016; J. A. C. Sterne et al., 2019).

The primary outcome of this study was to evaluate whether the use of technology-based applications increases the detection of melanoma in high-risk populations. Secondary outcomes include adverse events, safety data, and clinical implementation.

Search strategy

We searched PubMed, Medline (OVID) and Embase (OVID), Web of Science, and the Cochrane Library for articles published between January 1st, 2000, and July 16th, 2023, including the keywords "melanoma," "UV," "application," "artificial intelligence" (Complete search string can be found in the appendix). The time frame was limited to January 1st due to the limited availability of smartphones and apps in mHealth before this time. Duplicates were removed, titles and abstracts were reviewed independently, and full-text screening of all remaining articles was performed. In case of disagreement, a third researcher made the final decision. Reference lists of included articles were screened for further eligible articles, and a snowball search was conducted using ConnectedPapers (Ammar et al., 2018; Eitan, Smolyansky, Harpaz, & Perets). The entire screening process, including deduplication, was performed using the rayyan.ai software (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016).

Eligibility criteria

We included publications with the following criteria:
- Studies include computer-aided technologies (e.g., applications of artificial intelligence, smartphone applications, etc.) used to diagnose melanoma.
- Phase I, II III clinical trials, systematic reviews with meta-analysis, observational studies, technical notes with case description.
- Fitzgerald Skin Type I or II.
- Articles written in English, Spanish, Portuguese, German, French.
- Articles evaluating the following outcomes: safety, diagnosis, adverse events, and clinical implementation.

We excluded articles with the following criteria:
- Preclinical studies, conference abstracts, animal studies, ongoing trials with unpublished results, case reports, editorials, letters to the editor, technical notes without case descriptions, and narrative reviews.
- Pediatric patients (<18 years of age).
- Publications earlier than January 1st, 2000.
- Other interventions than computer-aided technologies.
- Adults with a history of premalignant or malignant skin lesions.
- Studies without information about population, intervention, comparator, and intervention.

Data extraction

Data regarding author, year of publication, number of patients included, safety outcomes, diagnostic outcomes, adverse events, detection rates, accuracy, and notions of clinical implementation were extracted and compiled for outcome analysis.

Results

Study Selection

We identified 819 articles, of which 472 remained after duplicate removal (Figure 1). After the screening, seven articles with 2694 patients remained in the outcome analysis (Börve et al., 2015; Börve, Terstappen, Sandberg, & Paoli, 2013; Hue et al., 2016; Lamel et al., 2012; Maier et al., 2015; Marek, Chu, Ming, Khan, & Kovarik, 2018; Massone et al., 2007). While reviewing our work, Philips et al. were identified as a further eligible study (Phillips et al., 2019). This article was previously excluded during title and abstract screening.

Study Characteristics

All articles covered the use of apps, with Maier

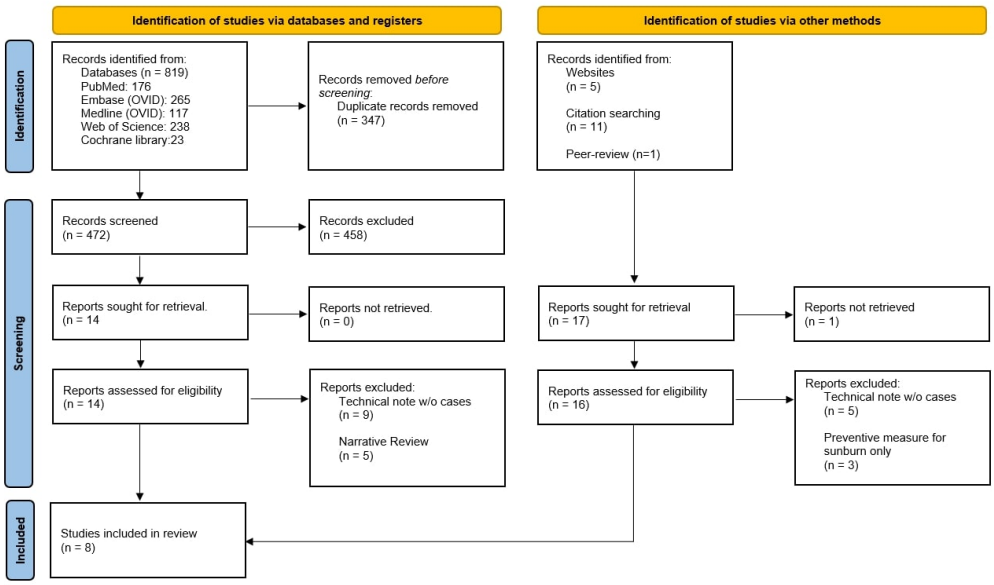

Figure 1: *PRISMA flow diagram of the literature review process.*

et al., Marek et al., and both studies of Börve et al. reporting apps free of charge. No AI, machine learning, or other computer-aided technologies were discovered in the primary search, yet Phillips et al. were added during the review (Phillips et al., 2019). We identified one RCT, one Pilot study, and five prospective cohorts. Diagnostics was the focus in all studies except for Marek et al., who focused on patient adherence. Melanoma was the exclusive disease under investigation in Maier et al. and Massone et al., while the remaining studies investigated multiple skin lesions alongside melanoma. Only Maier et al., Marek et al. and Börve et al. used different intervention and control groups while the remaining studies tested various interventions in the same population (Table 1) (Börve et al., 2015; Börve et al., 2013; Hue et al., 2016; Lamel et al., 2012; Maier et al., 2015; Marek et al., 2018; Massone et al., 2007). Philipps et al. analyzed an AI intervention trained to detect melanoma images taken by the smartphone of 541 patients in the UK (Phillips et al., 2019).

Primary outcome

Detection rates of melanoma in populations at risk were reported by Lamel et al. (100%) and Maier et al. (73%)(Lamel et al., 2012; Maier et al., 2015). The overall accuracy of the app was reported by Maier et al. (81%), Börve et al. (50.7%) and Börve et al. (42%) (Börve et al., 2015; Börve et al., 2013; Maier et al., 2015). If compared to dermatologists assessing the lesions as the standard of care, concordance rates were reported by Lamel et al. (62%), Massone et al. (89%), and Börve et al. (57%) (Börve et al., 2015; Lamel et al., 2012; Massone et al., 2007). Phillips et al. reported that their algorithm achieved an area under the receiving operator curve of 90.1% for all biopsied lesions, but results varied depending on the camera used (Phillips et al., 2019). No further comparison between different diagnostic tools was possible due to the limited number of studies reporting such outcomes.

Secondary outcomes

All articles report application in routine clinical practice; however, Lamel et al., Maier et al., and, post-hoc, Phillips et al. reported at least some limitation to the use of the application, while one study did not, and four did not mention potential limitations (Börve et al., 2015; Lamel et al., 2012; Phillips et al., 2019). There was no safety reporting, and consequently, no adverse events were reported.

Assessment of risk of bias in individual studies

Risk of bias in the RCT was found to be of some concern (Figure 2) while Risk of bias in the observational studies was found to be low in three studies (Hue et al., 2016; Lamel et al., 2012; Massone et al., 2007; Phillips et al., 2019) and moderate in the remaining studies (Börve et al., 2015; Börve et al., 2013; Maier et al., 2015).

Discussion

One of these is climate change, specifically the alteration in the temperature and ultraviolet radiation (Dzwierzynski, 2021; Parker, 2021). Exposure to UVR can be carcinogenic and is implicated as the primary cause of skin cancer due to DNA damage that promotes and initiates tumor formation. The interaction of UVR with increasing temperatures may amplify cutaneous photocarcinogenesis (Parker, 2021). Malignant melanoma is relatively uncommon in comparison with other types of skin cancer, but its incidence has increased as well. If one has sufficient indicators to suspect skin cancer, a cascade of examinations follows. Further diagnostic tools are warranted to facilitate diagnostics and reduce time loss while preventing unnecessary visits due to regular lesions. Different tools have been proposed, including vascular dermoscopy, reflectance confocal microscopy, teledermatology, computer-aided diagnostic techniques, and smartphone application (Phillips et al., 2019). Based on this premise, we conducted the present review and identified eight articles discussing some form of computer-aided technology directly compared to the current standard treatment. Although the quality of most studies was moderate, two RCTs, one pilot study, and five observational studies could be identified. As far as it was assessable, the detection rate was 86.5±19.1 %; accuracy was 57.9±20.5%, and concordance to standard care was 69.3±17.2%, respectively. No safety events were reported. However, three studies reported limitations in clinical use.

The definite advantage of computer-aided technology is the relative fastness of diagnosis, ideally markedly earlier than with the current standard of care, and the potential widespread use of such tools among family physicians (Freeman et al., 2020). Since most apps provide various mechanisms to assess skin lesions, ranging from simple image transmission to dermatologists to advanced algorithms assessing risk scores of specific lesions, many patients but also primary care providers are inclined to use it (Freeman et al., 2020). With the widespread distribution of smartphones and the increasing image resolution quality of modern mobile phone cameras, these tools have a massive potential to shorten the time to primary diagnosis while simultaneously preventing unnecessary visits to primary care providers (Kassianos, Emery, Murchie, & Walter, 2015). Also, handling such tools becomes more straightforward in a progressively more "digital" society. Hue et al. could show that FTF examinations can be reduced to up to 53% (Hue et al., 2016). The more efficient management of those patients was especially beautifully shown by Börve et al. (Börve et al., 2015). However, several limitations, most prominently claims of lacking validity, erroneous referrals, and overestimated accuracy rates, surrounding these tools still exist (Buechi et al., 2017; Freeman et al., 2020; Kassianos et al., 2015). A review recently found that most applications are soon retracted from the market due to significant variances in diagnostic accuracy. However, some, namely SkinVision, also improved over time (Freeman et al., 2020). Another notable limitation is the target population, below the average age for skin cancer development; therefore, computer-aided tools might target the wrong population (Freeman et al., 2020). Interestingly, although the camera resolution of smartphones has increased over the years, their use in suboptimal conditions counters these advances and leads to poorer image quality, given by the often-complicated instructions of present computer-aided applications (Freeman et al., 2020). Such limitations were confirmed by Maier et al., who reported difficulties with incomplete, i.e., images without surrounding skin were not recognized, erosive lesions were falsely interpreted, and difficulty in mottled or darker skin tones was observed (Maier et al., 2015). Contrary to that, however, Börve reported excluding 0.4% (4/902) patients due to poor image quality, yet examinations were performed by the general practitioner (Börve et al., 2015). Unfortunately, erroneous transmission problems can occur in particular, but not all, computer-aided diagnostic tools (Lamel et al., 2012).

The main question of this work was to identify whether the sensitivity, specificity, and/or accuracy of diagnosis of computer-aided applications was superior to the current standard of care. Maier et al., who used the SkinVision application, reported high sensitivity and specificity rates of the application with 73% and 83%, respectively, which was, however, markedly worse than the clinical diagnosis by the dermatologist (sensitivity of 88% and specificity of 97%) (Maier et al., 2015). This agrees with the most recent literature, which found a sensitivity of 88% with a specificity of 79%; however, it is limited to one application, SkinVision (Freeman et al., 2020). Lamel et al. found that applying the computer-aided program was relatively simple and that agreement between FTF and teledermatologist was excellent with Cohen's Kappa of 82% (Lamel et al., 2012). Likewise, Börve et al. report higher accuracy in diagnosis in FTF visits, albeit computer-aided technologies are more accessible in primary triage Fields(Börve et al., 2015).

Limitations

Although this review was performed according to the PRISMA guidelines, several limitations must be

Author	Year	Desing	Number of participants (I/C)	Intervention	Intervention frre of charge (yes/no)	Detection rate of skin cancer (%)	Accuracy of intervention (%)	Concordance of intervention with SOC (%)
Lamel et al.	2012	pCS	86/86	App	No	100	NA	62
Maier et al.	2015	pCS	144/144	App	Yes	73	81	NA
Marek et al.	2018	RCT	18/17	App	Yes	NA	NA	NA
Hue et al.	2015	PT	289/289	App	No	NA	NA	NA
Massone et al.	2007	pCS	16/16	App	No	NA	NA	89
Börve et al.	2015	pCS	69/69	App	Yes	NA	50.7	57
Börve et al.	2015	pCS	772/746	App	Yes	NA	42	NA
Phillips et al. (post-hoc)	2019	pCS	541/541	AI	No	100	NA	90.1

pCS: prospective cohort study; **RCT:** randomized controlled trial; **SOC:** standard of care; **NA:** not applicable; **AI:** articifical intelligence

Table 1: *Study characteristics of the included studies.*

mentioned. First, the quality of the included studies was moderate, leaving room for improvement and heralding care in interpreting the outcomes proposed in this work. A literature search was only conducted within a limited number of databases and languages included, potentially causing relevant manuscripts to be missed. Second, no proper outcome analysis in the sense of a structured meta-analysis with pooled outcome measures was possible given the few studies identified and the sparsity of reporting the primary outcome of this systematic review, making these results only partly recommendable for clinical application. Most of the included studies were non-randomized trials, making selection bias and confounding a reality to keep in mind when interpreting the results of the current work.

Furthermore, the significant heterogeneity and limited number of included studies strongly indicate publication bias in the field of interest. Also, studies reporting AI were not found, limiting recommendations and conclusions for this subsection of mHealth. Two authors that have contributed significantly to the present understanding of skin cancer detection using mHealth were Maier et al. and Freeman et al., both currently serving as consultants to the respective application employed, heralding care in interpreting the results of their studies. Lastly, since we included language and time restrictions in our selection criteria, this study risks not including a complete account of all available literature on the topic.

Conclusion

Although current studies show a promising future exclusively for health apps to detect melanoma in high-risk populations with excellent detection, accuracy, and agreement rates, albeit below the standard of care ranges, further studies should focus on improving the factors above with more extended follow-up periods and more diverse populations for adequate implementation in clinical practice. In contrast, patients and healthcare providers must know these applications' limitations. Randomized multicenter placebo-controlled trials with a larger sample size are required to measure the effectiveness and compare it with the SOC.

Author Contributions

Literature search: TJH; Article screening: TJH, HO, JM, SAB, CC, AVV, JC, KM, EM, KD, FMGS, LL, MR, TCR, DR, APV, AA, HB, EM, MCP, AG, AA, LGDJ; Data extraction: CC, TJH, APV, FMGS, MR, DR; Manuscript drafting: TJH, HO, JM, SAB, CC, AVV, JC, KM, EM, KD, FMGS, LL, MR, TCR, DR, APV, AA, HB, EM, MCP, AG, AA, LGDJ; Critically reviewing a manuscript: TJH, HO, JM, SAB, CC, AVV, JC, KM, EM, KD, FMGS, LL, MR, TCR, DR, APV, AA, HB, EM, MCP, AG, AA, LGDJ; Permission for publication: TJH, HO, JM, SAB, CC, AVV, JC, KM, EM, KD, FMGS, LL, MR, TCR, DR, APV, AA, HB, EM, MCP, AG, AA, LGDJ.

Funding

This research received no external funding.

Conflicts of Interest

TJH receives a personal scholarship jointly by the Swiss Academy of Medical Sciences and the Swiss National Science Foundation. The remaining authors report no conflict of interest.

References

- Ammar, W., Groeneveld, D., Bhagavatula, C., Beltagy, I., Crawford, M., Downey, D., . . . Etzioni, O. (2018). Construction of the Literature Graph in Semantic Scholar. Paper presented at the North American Chapter of the Association for Computational Linguistics.
- Börve, A., Dahlén Gyllencreutz, J., Terstappen, K., Johansson Backman, E., Aldenbratt, A., Danielsson, M., . . . Paoli, J. (2015). Smartphone teledermoscopy referrals: a novel process for improved triage of skin cancer patients. Acta Derm Venereol, 95(2), 186-190. doi:10.2340/00015555-1906
- Börve, A., Terstappen, K., Sandberg, C., & Paoli, J. (2013). Mobile teledermoscopy-there's an app for that! Dermatol Pract Concept, 3(2), 41-48. doi:10.5826/dpc.0302a05
- Buechi, R., Faes, L., Bachmann, L. M., Thiel, M. A., Bodmer, N. S., Schmid, M. K., . . . Lienhard, K. R. (2017). Evidence assessing the diagnostic performance of medical smartphone apps: a systematic review and exploratory meta-analysis. BMJ Open, 7(12), e018280. doi:10.1136/bmjopen-2017-018280
- Clark, J. M., Sanders, S., Carter, M., Honeyman, D., Cleo, G., Auld, Y., . . . Beller, E. (2020). Improving the translation of search strategies using the Polyglot Search Translator: a randomized controlled trial. J Med Libr Assoc, 108(2), 195-207. doi:10.5195/jmla.2020.834
- Dzwierzynski, W. W. (2021). Melanoma Risk Factors and Prevention. Clin Plast Surg, 48(4), 543-550. doi:10.1016/j.cps.2021.05.001
- Eitan, A., Smolyansky, E., Harpaz, I., & Perets, S. ConnectedPapers. Retrieved from https://www.connectedpapers.com/
- Freeman, K., Dinnes, J., Chuchu, N., Takwoingi, Y., Bayliss, S. E., Matin, R. N., . . . Deeks, J. J. (2020). Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. Bmj, 368, m127. doi:10.1136/bmj.m127
- Hue, L., Makhloufi, S., Sall N'Diaye, P., Blanchet-Bardon, C., Sulimovic, L., Pomykala, F., . . . Crémieux, A. C. (2016). Real-time mobile teledermoscopy for skin cancer screening targeting an agricultural population: an experiment on 289 patients in France. J Eur Acad Dermatol Venereol, 30(1), 20-24. doi:10.1111/jdv.13404
- Kassianos, A. P., Emery, J. D., Murchie, P., & Walter, F. M. (2015). Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. British Journal of Dermatology, 172(6), 1507-1518. doi:10.1111/bjd.13665
- Lamel, S. A., Haldeman, K. M., Ely, H., Kovarik, C. L., Pak, H., & Armstrong, A. W. (2012). Application of mobile teledermatology for skin cancer screening. J Am Acad Dermatol, 67(4), 576-581. doi:10.1016/j.jaad.2011.11.957
- Maier, T., Kulichova, D., Schotten, K., Astrid, R., Ruzicka, T., Berking, C., & Udrea, A. (2015). Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result. J Eur Acad Dermatol Venereol, 29(4), 663-667. doi:10.1111/jdv.12648
- Marek, A. J., Chu, E. Y., Ming, M. E., Khan, Z. A., & Kovarik, C. L. (2018). Piloting the Use of Smartphones, Reminders, and Accountability Partners to Promote Skin Self-Examinations in Patients with Total Body Photography: A Randomized Controlled Trial. American Journal of Clinical Dermatology, 19(5), 779-785. doi:10.1007/s40257-018-0372-7
- Massone, C., Hofmann-Wellenhof, R., Ahlgrimm-Siess, V., Gabler, G., Ebner, C., & Soyer, H. P. (2007). Melanoma screening with cellular phones. PLoS One, 2(5), e483. doi:10.1371/journal.pone.0000483
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. Syst Rev, 5(1), 210. doi:10.1186/s13643-016-0384-4
- Parker, E. R. (2021). The influence of climate change on skin cancer incidence - A review of the evidence. Int J Womens Dermatol, 7(1), 17-27. doi:10.1016/j.ijwd.2020.07.003
- Phillips, M., Marsden, H., Jaffe, W., Matin, R. N., Wali, G. N., Greenhalgh, J., . . . Palamaras, I. (2019). Assessment of Accuracy of an Artificial Intelligence Algorithm to Detect Melanoma in Images of Skin Lesions. JAMA Netw Open, 2(10), e1913436. doi:10.1001/jamanetworkopen.2019.13436
- Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., . . . Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. Bmj, 355, i4919. doi:10.1136/bmj.i4919
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., . . . Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. Bmj, 366, l4898. doi:10.1136/bmj.l4898
- WHO. (2022). Skin Cancer. Retrieved from https://www.iarc.who.int/cancer-type/skin-cancer/