

Peer-review Comments and Author Responses

Reviewer 1

1. *Although I understand there must be almost no publications in this area before January 1st, 2000, why was the search limited to this time-frame?*

Thank you for this excellent observation. Our rationale for the limitation to searches after January 1st 2000 was, that artificial intelligence and smartphone applications in the sense of mobile health for the broad population occurred primarily after 2000 with the introduction of the first smartphones. Thus, the search was limited to this time frame. We added a sentence describing this circumstance.

2. *Since you're conducting a systematic review, is it appropriate to also include previous systematic reviews in your eligibility criteria? If there were previous systematic reviews conducted on this specific topic, why is your systematic review needed? How is it novel?*

Many thanks for this important comment. We searched for further reviews mostly because we wanted to make sure that we have all available literature included as well as potential meta-analysis covered. As such, we also wanted to see, whether new information was available since the last systematic review leading to our review being the currently most comprehensive approach to the field.

3. *What precisely is the definition of "adults" and "pediatric" patients in your study?*

Thank you for this observation. We defined adults as participants that have reached or surpassed their 18th birthday. All participants below that cut-off were considered paediatric patients. We added a respective statement.

4. *The title of the systematic review describes "Does the use of an artificial intelligence or a smartphone application increase the **detection** of melanoma in adults at high risk for development of melanoma". During the description of the eligibility criteria, the outcomes described for analysis are: "safety, diagnosis, adverse events, clinical implementation". Does the title reflect the breadth of the review?*

Thank you for this excellent comment. As we have specified, the primary outcome of this manuscript was to see whether technology-based applications increase the detection of melanoma. As such, the title reflects the most important component of our work, while secondary outcomes as safety events and clinical implementation support the primary outcome but, in our opinion, do not necessarily mean to be included in the title.

5. *How did the reviewers identify and remove duplicate articles before screening? Did this involve the usage of computer software? If so, which one?*

The whole search and data acquisition process was performed using the software rayyan.ai which has an automatic deduplication mechanism built in. We added a respective sentence to the manuscript.

6. *The authors identify only 8 articles which fit inclusion criteria, as the criteria did not exclude reviews, why wasn't Kassianos et al. 2015 (DOI: 10.1111/bjd.13665) included?*

Many thanks for this excellent observation. However, Kassianos et al. is a narrative review, therefore not meeting the eligibility criteria and thus not being included.

7. *In the construction of the search string, why was the "(“ultra-violet ray*”[tiab] OR “ultra violet ray*”[tiab] OR “ultraviolet ray*”[tiab] OR “UV ray*”[tiab] OR “ultra-violet light”[tiab] OR “ultra violet light”[tiab] OR “ultraviolet light”[tiab] OR “ultraviolet radiation”[tiab] OR “ultra-violet radiation”[tiab] OR “ultra violet radiation”[tiab] OR “UV radiation”[tiab] OR “ultraviolet rays”[Mesh])" string included?*

UV term in search string: Many thanks for this question. As UV rays are the most important risk factor for the development of skin cancer, it was only natural to include this term in the search string as there might have been apps specifically assessing UV levels. As such, it was necessary to search for UV related terms for a wholistic search.

8. *In the exclusion criteria, no mention was made of exclusion of papers which are mainly considering apps that use artificial intelligence, later on, the manuscript says: "All articles covered the use of apps (free-of-charge at 62.5%) while neither AI nor machine learning or other computer-aided technologies were covered." There are, however, papers that could fit inclusion criteria and deal with artificial intelligence such as Phillips et al. 2019 (DOI: 10.1001/jamanetworkopen.2019.13436).*

Many thanks for this keen observation. Phillips et al. like many other AI training papers used online available dermatologic image data banks. As such, most of the AI based papers do not qualify for inclusion in our study as they do not meet all eligibility criteria why it was most likely ruled out in the title and abstract screening process. We reconsidered and post-hoc included Phillips et al in our review.

9. *In study characteristics it is described "while one study looked exclusively at sunburn". In eligibility criteria, the authors defined "Studies including computer-aided technologies (e.g., applications of artificial intelligence, smartphone*

applications etc.) used for diagnosis of melanoma". Does the cited study fit inclusion criteria? Specifically, which of the included studies only analysed sun-burns?

Thank you for that excellent observation, we totally agree, this study must be excluded as it somehow passed our screening process. We have removed the publication by Hacker et al. from our analysis.

10. *I did not find a table describing the application of ROBINS to the non-randomized studies.*

We have added the ROBINS table to the supplemental material.

11. *As the current work does not include a meta-analysis, wouldn't it be better to report the results of the papers individually?*

Thank you for this comment. We agree and have modified the results section accordingly.

12. *Inclusion of a funnel plot could be useful in assessing for publication bias.*

Thank you for this comment. However, due to the heterogeneity in the results reported by the different publications and the small number of trials included we refrained from reporting a funnel plot as there is clearly underreporting of this subject. We have added a sentence in the limitation section addressing this point.

13. *I'd consider limiting the outcomes analysed and the types of study that are included in the review.*

Thank you for this suggestion. As the outcomes and study types were defined before the review was conducted, removing either types of studies included or change outcomes defined would result in a protocol violation. Therefore, we would refrain from limiting outcomes and study types included.

Reviewer 2

14. *While the introduction provides a thorough background on the significance of melanoma detection and the use of technology-based applications. However, it would be helpful to emphasize the importance and distinctiveness of your study, stating why this research is essential, what it would add to the literature, and how it addresses a gap in the current literature, which would strengthen the introduction.*

Many thanks for this insightful comment. We have adapted the introduction accordingly.

15. *This review mentions biases and limitations very frequently, some concerning ones are; the exclusion of some language publications which may lead to language bias, the restriction of the publication date range which might result in the omission of significant historical studies, and the small number studies included from a large pool of 400+ initial studies. I suggest the following: Re-evaluate your inclusion/exclusion criteria, re-run the screening processes and/or use a different tool to screen, and, to enhance this review's value and relevance, it would be beneficial to discuss potential strategies or actions taken to mitigate these biases and limitations, underscoring why this review could contribute significantly to the existing literature. Finally, if after these steps more studies were included, consider running a formally structured meta-analysis.*

Thank you for this important comment. We have explained our rationale for limiting the time frame of the search in the comments for reviewer #1, answer #1. Regarding language restrictions, we have limited our search to languages spoken by the members of the study team as inclusion of more languages would have been associated with disproportional effort to organize translations or translators, we have refrained from this step. We have added a sentence addressing the limitations mentioned by you but would refrain from rerunning the search as changing the search criteria would mean violating our study protocol.

16. *The review highlights that the included studies primarily concentrated on one specific type of technology application, namely smartphone apps, with no mention of artificial intelligence or machine learning. This limitation should be discussed in more depth as the study aimed to assess different types of technology applications. The absence of diversity in the types of technology-based applications explored may limit the comprehensiveness of this review. Alternatively, the title and aim of the review could be modified to focus exclusively on apps, or the review could assess and specify whether these mobile apps incorporated artificial intelligence.*

Thank you for this excellent comment. Please refer to comment #9 for reviewer #1. We addressed this in the limitations section.

17. *The discussion section focuses on the failures encountered while using those apps. Consider adding a paragraph on the implications of the study for future research. What are the key areas that require further investigation? Based on the studies that were reviewed, what could be investigated in the future to enhance the use of apps and AI in the diagnosis of melanoma? What do the authors suggest/recommend?*

Many thanks for this valuable comment, we totally agree. We have added the respective section and recommendation to the limitations.

18. *Freeman et al. and Maier et al. seem to have been referenced extensively in the discussion section. Considering the potential conflicts of interest among the authors of these papers, it is advisable to conduct a more thorough assessment to determine whether any of the authors had direct affiliations or relationships with SkinVision. I observed that three of the authors of the study by Maier et al. have served or are currently serving as consultants for SkinVision.*

Thank you for this excellent observation, you are right, and we completely agree. We have added a respective sentence to the discussion.

Reviewer 3

19. *Your introduction is very well written and gives a general background of what is known in this area. However, when you present your research question, it is somewhat generic: "...the purpose of conducting this study was to systematically review the current knowledge available regarding the use of technology-based applications,". Here, I think the aim should be more specific, such as the question posed in the title: "Does the use of an artificial intelligence or a smartphone application increase the detection of melanoma in adults at high risk for development of melanoma: a systematic review". So the aim of the study, according to the title, should be: "can algorithm based (artificial intelligence) smartphone apps increase detection of melanoma". It seems that what you wanted to study was this, which is a very relevant and interesting question. According to this question, we would go directly to Freeman, et al, 2020, which is a systematic review of precisely the question you have as your aim of study. The study aim in Freeman's review is stated as: "In our review, we aim to report on the scope, findings, and validity of the evidence in studies that examine the accuracy of all apps that use inbuilt algorithms to identify skin cancer in users of smartphones". As the review was published in 2020, you would have to add any new study regarding algorithm-based smartphone app for detecting melanoma, and perform a new systematic review adding the new papers, giving a substantial contribution in advancing scientific knowledge. I agree that this would be an incredible undertaking since you would have to re-write almost the whole manuscript. That is...unless you change your title to something like: "The role of smartphone-based apps in aiding screening, faster diagnosis, and prevention in skin cancer. A systematic review." I think then the title will be more consistent with what you have written and the aim of the study will go in the same direction.*

Thank you for this very well formulated observation. We agree that we would need to change the title to reflect more precisely what our review describes and have made the corresponding adaptations.

20. *Reading the abstract, the last phrase of conclusion section (page 2, line 16) is quite out of place: the study you performed does not allow you to conclude that "Further studies in more diverse populations assessing effectiveness as well as improvement of*

current application algorithms are warranted.” Nowhere in your study you mention that “diverse populations” is an issue to increase detection of melanoma. It could be true, but your study does not support that conclusion. Therefore, I would remove that sentence altogether.

Thank you for this comment. We agree and removed the respective sentence.

21. *I unfortunately did not see a table with the characteristics of the included studies, which is necessary (design of the study, sample size, main findings). There is very little information about the studies you included. The same problem arises when you describe the study characteristics (page 5, line 6). In this paragraph, you describe the studies very briefly and do not specify which study is the one you are describing (you just put the 8 study’s reference at the end of the paragraph). In this way, the reader does not know which study you are referring to, for example, when you say: “Melanoma was the exclusive disease under investigation in two (25%) studies while one study looked exclusively at sunburn...” So, which two studies examined only melanoma as the exclusive disease under investigation? You need to put the reference after each study description so the reader can identify which paper you are talking about.*

Thank you for this important observation, we have added a table 1 with the respective study characteristics and references.

22. *When you describe the primary outcome (page 5, line 19), I could not understand how you calculated the mean cumulative detection rate by apps: “...Detection rates of melanoma in populations at risk were reported by two studies with a mean cumulative detection rate by apps of 86.5±19.1 % compared to the standard of care (Lamel et al., 2012; Maier et al., 2015)”. Because in the abstract you mention that this detection rate was “across all included studies”(page 2, line 13). To make matters a little worse, when you mention detection rates what are you comparing with? You see, ... I am afraid that you cannot make a general conclusion of detection rates because only the study by Maier et al., 2015 is the one that compares an artificial intelligence app detection of skin cancer, comparing it to a dermatologist, with histology as the gold standard. The other studies use apps for a variety of reasons (to aid in screening, for prevention, etc.) such as to take pictures of lesions and see if an online dermatologist can classify the lesion as malignant, compared to a face to face in-person consultation by dermatologist. So, in this last case, the app is simply taking pictures of the lesion to be analyzed by a dermatologist (the app is not detecting cancer based on artificial intelligence). Therefore, I think you should describe each study individually, since pooling studies is not possible due to the nature of the selected studies. Again, a table with a summary of each study would be very valuable.*

Many thanks for this important observation. Please see comment #12 for reviewer #1, where we describe that we have adapted the results accordingly.

23. *The discussion is very well written and by the end of it the reader captures the overall message that smart-based app technology can be helpful in many ways, but still needs improvement in artificial intelligence algorithms to be useful in making detection of skin cancer faster and earlier, to impact patient care. However, smartphone based apps can have other roles as clearly expressed in the conclusion remarks made by Freeman, et al 2020, and which are in agreement with your review: "Smartphones and dedicated skin cancer apps can have other roles; for example, assisting in skin self-examination, tracking the evolution of suspicious lesions in people more at risk of developing skin cancer,47 48 or when used for store and forward teledermatology.49 50 However, healthcare professionals who work in primary and secondary care need to be aware of the limitations of algorithm based apps to reliably identify melanomas, and should inform potential smartphone app users about these limitations." I would suggest you to end the discussion with some similar remarks.*

Many thanks for this comment. We have added a respective statement as end of the discussion.

Reviewer 4

24. *The title could be shorter, a suggestion is "Computer-aided technologies for melanoma detection in high-risk adults: a systematic review. If there were zero articles covering artificial intelligence why should that be included in the title?"*

We would like to our response for reviewer #3 comment #1 where we address the title. We have adapted the title accordingly.

25. *Introduction: Is this prevalence or incidence? If you are referring to incidence, it should be: "325,000 new patients in 2020"*

Thank you for this observation, the number refers to the incidence and we have added the correction accordingly.

26. *Introduction: Is it ambiguous or unknown? (Paragraph 3)*

This specific statement refers to the ambiguous rather than to the unknown.

27. *I suggest including the PICOS for this systematic review, the target population is missing in this section. If the population is adults at high risk of melanoma that should be stated here and specify what you consider as high risk.*

Many thanks for this important remark. We have added the PICOT scheme to the search strategy in the supplemental material.

28. *How did you perform that screening process? It could be more detailed.*

Many thanks for this comment. We would like to draw the attention of the reviewer to the methods section detailing this process as well as to figure 1 showcasing the steps undertaken.

29. *I suggest the use of a table to summarize the characteristics and findings of the included studies.*

We refer to our comments to reviewer 3.

Reviewer 5

We would like to thank reviewer 5 for his kind review of our work. We have implemented the suggested changes in our manuscript.

Reviewer 6

30. *I strongly suggest double checking the RoB chart (Figure 2). There are 8 articles which were included in the final review. However, Figure 2 shows 2 studies only.*

Risk of Bias chart: Many thanks for your attentive comment. The RoB 2 tool is to be used solely for RCTs that is why only two studies are included in the respective chart. As addressed in comment #1 for reviewer 1, the remaining studies are depicted in the ROBINS tool, which we added to the supplemental material.

31. *Also, you mentioned that "a third researcher (AA) made the final decision." What does AA mean? It does not appear in the abbreviation section.*

Third reviewer: Many thanks for this comment. The AA refers to the author Ashraf Ahmed. All further authors are referred to in the Authorship contribution.

32. *I would also suggest explaining the limitation section more in detail. Finally, consider adding a table 1 (characteristics of the included studies).*

Table 1: Many tanks for this comment. We addressed this issue in comments #3 and #6 for reviewers 3 and 4 respectively.