Reviewer's report

Title: Smartphone apps for calculating insulin dose: a systematic assessment

Version: 1 Date: 12 January 2015

Reviewer: Urs-Vito Albrecht

Reviewer's report:

This paper is a systematic review and assessment of various quality related aspects of mobile apps that can be used for calculating medication doses, or, more specifically, insulin calculators for patients with diabetes. With their study, the authors address an important question for apps that are to be used in this context: Do existing apps contain adequate and accurate information and algorithms enabling patients with diabetes to safely calculate the correct dosage for insulin that is necessary for their current situation (based on data entered by the patient)? In their evaluation, the authors have included various factors that may potentially lead to problems, ranging from data entry issues to ambiguous terminology being used or incorrect or unverifiable calculation methods.

The paper is clearly structured and well written. The selection and assessment processes used within the study are listed in sufficient detail and the conclusions. However, there are a few very minor issues I would like to mention:

Discretionary Revisions

Abstract: Within the “Methods” subsection, the authors state that they performed a “Systematic assessment of all English-language rapid/short-acting insulin dose calculators available for iOS and Android”, whereas in the “Results” subsection, they specify that “Included apps had been downloaded at least 105,000 times, and likely substantially more.”. Without having read the rest of the paper, it is not immediately clear whether “downloaded at least 105,000 times” (which is hard to confirm in the case of iOS based app since Apple does not provide download numbers) was in fact by itself an exclusion criterion or whether this was simply a result of the other exclusion criteria that were applied. Since in its current form this statement does not really provide any essential insights for readers of the abstract, I would suggest to simply remove it here. On the other hand, I would add a statement within the results listed in abstract that is currently only found in the full version of the results: “Only one app, for iOS, was issue-free according to our criteria.” as this will certainly pique the interest of readers.

Abstract: While certainly a matter of taste, within the results subsection of the abstract, I would suggest to include not only the percentages, but also the absolute numbers for the various problems that were identified with respect to the apps. On a similar note, this should also be done throughout the paper in places where currently, only percentage values are given.

Methods: Since the review was performed by at least two investigators, it might be helpful to provide readers with information about their professional
backgrounds to give them a better idea about whether the reviewers are able to judge the selected apps from a technical, medical or layperson’s point of view.

Results: For apps that were available for both Android and iOS, were relevant differences identified regarding functionality?

Results: The authors state that the clinical disclaimers were analysed. According to which criteria was the analysis performed?

Results: The authors state that for a small number of apps, developers dissuade use for treatment or other medical purposes. It would be helpful to discuss why those apps were not to be excluded from the overall analysis and whether including these apps could have biased the results.

Discussion: Limitations of the study should be discussed more extensively.

Discussion: While I agree with the statements the authors provide with respect to regulatory issues, app accreditation programs and so on and believe that regulation of apps such as the insulin dosage calculators that were evaluated in the presented study (and have the potential to cause serious harm for a patient) is important, I do not believe that it can be sufficiently enforced in all cases since, as the authors recognise, many medical apps and health apps provided within official app stores are personal projects of medical laypersons and "Personal projects are unlikely to have the resources available to pay for expert guidance, formal testing, or updates as new evidence emerges.". As such, many of them will also lack resources that enable them to undergo regulatory processes or to obtain other certification. In contrast to the authors, while it is certainly true that "Coordinated programs of surveillance for harms resulting from medical app use, whether run by regulators, accreditation programs or researchers, should be a priority.", I wonder whether it is sufficient to rely on such measures. Rather, I believe that educating clinicians as well as patients about potential problems of using medical apps and health apps is even more important in order to enable them to assess the most important aspects of an app before they decide to use it. This also includes providing them with standardized tools that enable them to make such a decision, e.g. in the form of an easy to use checklist.

Discussion: Since the authors present their methodology as a template for reviewing apps, they should also discuss their work in the context of other review schemes and standard reporting tools for apps. How does their work relate to such review schemes and tools?

Conclusion: The authors stated that “the majority of apps in this study were packaged professionally, with consistent layout and design”. Is this meant in the context of usability? If so, I doubt the validity of this statement since only two reviewers performed the evaluation. Also, if usability testing was done, which methods were used? Were standardized tools employed? I was unable to find any hints regarding this.

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:

I declare that I have no competing interests