Dr. Spencer Smith, Associate Professor
Computing and Software Department
Faculty of Engineering
McMaster University
1280 Main Street West
Hamilton, Ontario, Canada L8S 4K1

Telephone: (905) 525-9140 ext. 27929
email: smiths@mcmaster.ca
Web: http://www.cas.mcmaster.ca/~smiths

January 25, 2025

Dr. Anthony McGoron
Editor-in-Chief
Critical Reviews in Biomedical Engineering

Re: Revisions to CRB-57015, State of the Practice for Medical Imaging Software Based on Open Source Repositories

Dear Dr. McGoron:

Thank you and the reviewers for the feedback on our submission. The reviewers provided thoughtful and constructive comments. In response to your e-mail, dated January 3, 2025, we have revised the paper to incorporate the requested revisions. We provide a summary below. The revised submission also includes a "diff" version of the paper showing the additions and deletions.

Some of the reviewer's questions are best answered by citing our related work. Therefore, the new version undoes the redaction originally added for the double blind review.

Reviewer #1

1. The relative priorities of each software quality were set equal. Couldn't some different weighting have been considered? Perhaps one that comes from interviews with developers or cited in papers?

   The reviewer is correct that other weightings could be considered. We left this out of the scope of the original submission because of space considerations and because determining a meaningful alternative weighting is difficult. Although Nguyen-Hoan et al. (2010) (A Survey of Scientific Software Development) provides survey data on the relative importance of various software qualities, their list of qualities differs from ours. We are not aware of a recent survey that assesses the relative importance of qualities for scientific software. Moreover, even if we have quality priorities for all scientific software, our paper focuses only on medical imaging software. The priorities for medical imaging software likely differ from the priorities for other scientific applications. Furthermore, the priorities likely differ between different medical imaging applications. Rather than speculate on the relative priorities, we used the simplest case of equal priority. For

any readers that would like to apply their own priorities, we can direct them to the spreadsheet used to generate the results shown in the paper. With some simple modifications the plots can regenerated with whatever priority ranking a user wishes to see. To make this clear, we have added the following text to the revised paper:

(Alternative priority schemes can be investigated by changing the Prioritization Matrix in AHP_Template.xlsx available on GitHub.)

2. In addition to github, could any other reference have been used, such as the number of citations in papers or relevant websites in the field?

   Although the idea of including citations in papers is a good one, we have excluded this data because we feel it is unreliable. Unfortunately, software is infrequently and inconsistently cited even when it is used (as observed by Smith et al (2016) (Software Citation Principles)). Similar inconsistency problems exist for data on websites. For other readers who have the same observation as the reviewer, we have added the following text to the revised version:

   We considered also comparing software citations, but this measure of popularity would be unreliable since software is infrequently and inconsistently cited (SmithEtAl2016).

3. Where are the quality ratings for each software?

   The quality ratings for each software are summarized in Figure 2 (Overall AHP scores with an equal weighting for all 9 software qualities). To keep the length of the paper reasonable, we do not have figures showing the ranking of the software packages for each quality, but this information is available through the citations given in the now unredacted revised version of the paper. Some of the relevant citations include Dong (2021) (Assessing the state of the practice for medical imaging software, MEng thesis), and Smith et al (2024) (State of the practice for medical imaging software, extended report on arXiv).

4. Caption tables above them, not below them.

   This has been fixed in the revised manuscript.

5. What are the possible reasons for the differences between the ranking found and that of GitHub, in addition to the impossibility of installing some of them?

   Other possible reasons for the differences in ranking are covered in the original submitted paper in the following text: "We weighted all qualities equally, which is not the likely the same weighting that users implicitly use. To properly assess this would require a broad user study. Furthermore our measures of popularity (like stars) are only proxies *which are biased towards past rather than current preferences (Szulik2017), as these are monotonically increasing quantities. Finally there are often more factors than just quality that influence the popularity of products." Given that this wasn't obvious to the reviewer, we strengthened the topic sentence that introduced the reasons with the following text:*

Besides the installation problem, another possible reason for discrepancies between our ranking and GitHub popularity is that we weighted all qualities equally, ...

6. Some software could not be installed. What problems were found? Why not remove it from the list?

   In the interest of space, we have chosen to not include the details in on installation problems for GATE, dwv, and DICOM Viewer in the paper. Instead, the additional details are given in the extended version of the paper on arXiv (Smith et al (2024)) and in the associated Masters thesis (Dong (2021)). The software with installation problems was not removed from our review because we aimed for as complete an overview as possible. Moreover, the installation problems only affected the score on installability, reliability and robustness. The other qualities are based on the artifacts in the repository, not on running the software. To clarify this for readers of the paper, we have added text to the methodology for grading in the revised manuscript:

   We explicitly note any software that cannot be installed, but for completeness we still measure it for the remaining qualities, since only reliability and robustness require running the software.

Reviewer #2

1. The introduction section should contain the scope, significance of the research by summarizing current understanding and background information, stating the purpose of the work, and highlighting the potential outcomes. Also, there's no need to divide various subsection in the introduction.

   We believe that our introduction covers the scope (in Section 1.2 of the submitted version) the significance (in the first paragraph). We also believe that the current understanding, background information and potential outcomes are covered by the research questions. There is not a considerable amount of background information because there isn't a previous study (that we are aware of) that assesses the state of the practice for medical imaging software. The potential outcomes are the answers to the research questions.

   The criticism of the introduction, including the comment about not having subsections, is likely due to our covering the methodology in the introduction section. In the revised version, we have made the methodology its own section and removed the subsections from the introduction.

2. The title of the table should be placed above the table.

   This has been fixed in the revised manuscript.

3. It is not recommended to include tables in the conclusion.

   Although including a table in the conclusions is rare, we think it is appropriate for this paper. A table is a quickest way to summarize the software package quality. Summarizing the same

information in words would be hard to follow. A conclusion shouldn't introduce new information, but the table is not new information; it is an alternate view of the information presented in Section 4 of the paper. This new view makes some conclusions about the state of the practice clear, so it feels appropriate for the conclusion section.

Best regards,

Spencer Smith