*Response to reviewer comments on the manuscript “The importance of open science for biological assessment of aquatic environments” by M. W. Beck, C. O’Hara, J. S. Lowndes, R. D. Mazor, S. Theroux, D. J. Gillett, B. Lane, G. Gearheart. Submitted to PeerJ.*

*We thank the reviewers for providing detailed and thoughtful comments on our manuscript. Our responses to these comments are below. Line numbers refer to those in the original pdf submission.*

**Editor comments:**

A number of important considerations have been raised and need to be fully addressed in the next version. In particular, narrowing down the title of the review to 'aquatic biological assessment' (reviewers 1 and 2), re-tailoring the introduction and discussion sections to a more prominent mention of the failures in reliability that brought about the open science movement (reviewer 1) while also expending the background beyond the USA context (reviewer 2), and perhaps providing more emphasis on what needs to be done for environmental managers to effectively train and apply ecogenomics and R-based bio-assessments (reviewer 3).  
  
*Thank you for the summarizing the main points raised by the reviewers. We have addressed each of these concerns in our detailed responses below. In short, we have modified the title to better reflect the scope of the review, revised the content to better contextualize the open science movement and included additional international examples, and included some text regarding molecular approaches as currently used in bioassessment.*

**Reviewer 1:**

Basic reporting

This article presents a review of open science practices in ecological science. I found it to be well written and informative. I did however find the approach taken to be quite limited to ecological aquatic biological assessment. I understand within the field biological assessment is the correct term but given the scope of the journal, biological assessment could be expected to include other areas of biology such as medicine. Perhaps clear statement of ecological angle in the title would resolve this issue. I also thought the introduction would benefit from a more prominent mention of the failures in reliability that brought about the open science movement. In its current form, it seems to advocate standardisation as much as open practices (although I see the two are related). I also found the description of the principals of open science to be slightly narrower than I had expected (see below). I am not aware of a similar review having been recently published.

*We appreciate the helpful comments and have addressed your concerns below. In brief, we have amended the title of the manuscript to the following: “The importance of open science for biological assessment of aquatic environments”. Please note that reviewer two had a similar concern. We have also added content regarding pre-processing, as described in our response below.*

*For the introduction, the following was added to line 77: “Open science and its ideals originated partly due to failures of reproducibility and biases in the primary literature that were revealed as systematic concerns in research fields with immediate implications for human health (Makel et al.* [*2012*](file:///C:\proj\bioassess_opensci\manu_draft.docx#ref-Makel12)*, Franco et al.* [*2014*](file:///C:\proj\bioassess_opensci\manu_draft.docx#ref-Franco14)*). These ideas and the failures that they address have slowly permeated the ecological and environmental sciences…”*

*It is also worth mentioning that the expansion of this work to include medical or clinical research would be beyond our current scope that focuses on bioassessment in aquatic environments. We agree that open science benefits other fields, but there are already several relevant papers that cover this topic (e.g., Vasilevsky et al. 2013* [*https://doi.org/10.7717/peerj.148*](https://doi.org/10.7717/peerj.148)*, Allen and Mehler 2019* [*https://doi.org/10.1371/journal.pbio.3000246*](https://doi.org/10.1371/journal.pbio.3000246)*, Wallach et al. 2018* [*https://doi.org/10.1371/journal.pbio.2006930*](https://doi.org/10.1371/journal.pbio.2006930)*)*

Experimental design

As stated above I found the coverage to be somewhat limited to ecological methodologies, but this is probably acceptable if the scope is clearer at the outset. More importantly, a major aspect of open science practice more broadly is the adoption of study pre-registration and registered reports. I thought the absence of a discussion of the impact of these practices was an omission. I would also suggest that were the authors to include a discussion of these practices it would be advisable to enlist the aid of someone with first-hand experience of implementing them. Adopting these practices is a major cultural shift in open science, which I don’t think can easily be left out (including them would certainly alter figures 1 and 2). Perhaps the distinction between confirmatory and exploratory research is not as relevant here as in other disciplines (e.g. neuroscience) and environmental science does not rely on hypothesis testing as much, in which case might that be noted? I also thought broader discussion of open access platforms and repositories (such as the [osf.io](http://osf.io/)) would benefit the reader.

*We agree that pre-registration needs to be discussed in our review, specifically regarding its application to environmental science as compared to confirmatory or more conventional research. We have added the following content to line 246 and have also modified figure 2.*

*“**An important practice that is currently not used in bioassessment for project planning is study pre-registration. This is a relatively new addition to the philosophy of open science that allows a research team to define their study procedures, expected outcomes, and statistical analysis plans in advance of the actual study (Munafó et al.* [*2017*](file:///C:\proj\bioassess_opensci\manu_draft.docx#ref-Munafo17)*). Although the standard scientific method may seem to support such proactive practices, pre-registration is an explicit declaration to make the intentions of a study design clear to avoid publication bias where only positive outcomes are reported and to prevent an interpretive result where the researcher retrospectively defines study objectives after initial results are obtained if they do not agree with expectations. This latter issue is a serious concern where scientists use postdiction with significant hindsight bias in place of prediction and conventional hypothesis testing to inform scientific discovery (Nosek et al.* [*2018*](file:///C:\proj\bioassess_opensci\manu_draft.docx#ref-Nosek18)*). Pre-registration has been used extensively in clinical research (Dickersin and Rennie* [*2003*](file:///C:\proj\bioassess_opensci\manu_draft.docx#ref-Dickersin03)*), where outcomes often have immediate implications for human health and well-being. In contrast, bioassessment studies often focus on developing applied products, where conventional hypothesis testing is less a concern. However, pre-registration could be an important tool for the environmental sciences where an explicit declaration of study intent as being exploratory or applied could prevent postdiction or an otherwise misuse of study results after a project is completed. Existing venues that support pre-registration of studies across multiple disciplines could be used in bioassessment study planning (e.g.,* [*Open Science Framework*](http://osf.io/)*,* [*AsPredicted*](http://aspredicted.org/)*).”*

Validity of the findings

The conclusions appear to be well supported and clearly made. My concerns here are centred upon the apparent omission of important aspects to open science practices, mentioned above.

Comments for the Author

I found this to be a manuscript of good quality. I did, however, feel a number of important open science practices, partially pre-registration, where not discussed and perhaps should have been. Perhaps the authors could consider narrowing the scope of the article and discussing open data and materials practices, rather than open science?

*We have amended the title and added content regarding pre-registration. Please see the comments above.*

**Reviewer 2:**

Basic reporting

The English is clear, unambiguous, and professional.  
Although the Introduction and background show the context, there are some issues which need to be taken into account, including the absence of some relevant literature. Some comments to the Introduction are shown below.

- Line 1: I would add ‘aquatic’ in the title, before ‘biological assessment’, since the authors focus mainly in that and this should be highlighted since the very beginning

*The title has been revised to better reflect the scope of the review: “The importance of open science for biological assessment of aquatic environments”*

- Line 24: ‘Democratize science’, probably this is an opinion of the authors, and my response could be considered also as an opinion. As such, you can take it into account or not. I don’t like this expression, since it seems that the general scientific process is not democratic. However, in my opinion, is one of the most scrutinized human activities, since the moment in which scientists apply to public calls for projects and then publish the results either in open access or non-open access journals, since they are checked by editors and reviewers, trying to ensure replicability of the publications. Are open access principles more ‘democratic’? I disagree: they allow more interaction, more replicability, more transparent, more responses to new questions,… but I cannot see that as more ‘democratic’.

*We agree with this description and have replaced “democratize” with “improve”. Our original intent was to emphasize increased collaboration and openness among peers, which is not entirely described by “democratize”, as the reviewer notes.*

- Lines 42-46: it is good to have some examples of the legislation driving monitoring and assessment in these three areas. However, there are many other countries/regions (e.g. China, South Africa, Australia, New Zealand, etc.), as well as other legislation (e.g. Oceans Act in several countries, Marine Strategy Framework Directive, international conventions, etc.), which deserve also to be cited. As this could be too long, I suggest you adding a citation here. To my best knowledge, although it has more than 10 years, the most complete review in this specific topic is this paper, collating the legislation around the world in monitoring and assessment: Borja, A., S. B. Bricker, D. M. Dauer, N. T. Demetriades, J. G. Ferreira, A. T. Forbes, P. Hutchings, X. Jia, R. Kenchington, J. C. Marques, C. Zhu, 2008. Overview of integrative tools and methods in assessing ecological integrity in estuarine and coastal systems worldwide. Marine Pollution Bulletin, 56: 1519-1537.

*This sentence was revised, including adding the suggested citation: “Legal mandates to assess biological condition have stimulated the development of bioassessment programs and tools in the United States (Clean Water Act, CWA), Canada (Canada Waters Act), Europe (Water Framework Directive), China (Environmental Quality Standards for Surface Water), South Africa (National Water Act), and elsewhere (Borja et al. 2008).”*

- Lines 48-49: regarding marine waters, one of the co-authors of this manuscript has also a paper that can be cited here, since it constitutes a good review of the current methods availability: Borja, A., M. Elliott, J. H. Andersen, T. Berg, J. Carstensen, B. S. Halpern, A.-S. Heiskanen, S. Korpinen, J. S. S. Lowndes, G. Martin, N. Rodriguez-Ezpeleta, 2016. Overview of integrative assessment of marine systems: the Ecosystem Approach in practice. Frontiers in Marine Science, 3: doi: 10.3389/fmars.2016.00020.

*The citation was added.*

- Lines 72-85: regarding the use of open science in bioassessment, in relation to stakeholders, to my knowledge, this paper below is the single one dealing with that. The second one is dealing with the transparency issues. Both could be useful in this context.  
o Borja, A., J. M. Garmendia, I. Menchaca, A. Uriarte, Y. Sagarmínaga, 2019. Yes, We Can! Large-Scale Integrative Assessment of European Regional Seas, Using Open Access Databases. Frontiers in Marine Science, 6: 10.3389/fmars.2019.00019.  
o Essl, F., F. Courchamp, S. Dullinger, J. M. Jeschke, S. Schindler, 2020. Make Open Access Publishing Fair and Transparent! BioScience, 70: 201-204.

*We have added a sentence referencing Borja et al. 2019 as an example: “Limited examples have suggested that open access databases can be leveraged to develop bioassessment products that increase transparency among stakeholders (Borja et al. 2019).”*

*The second reference was also added to emphasize new modes of publishing (as an addition to the previous sentence starting on line 78: “Open science has also influenced how research workflows are conceptualized in other disciplines (e.g., archaeology, Marwick et al. 2016, behavioral ecology, Ihle et al. 2017, hydrology, Slater et al. 2019, vegetation sciences, Collins 2016) and has enabled in a shift towards publishing structures that are more fair and transparent through open access (van Oudenhoven et al. 2016; Essl et al. 2020).”*

The structure is clear, and of broad and cross-disciplinary interest.  
I don’t know any other recent review. However, one of the co-authors of this paper, is the first author or another paper on the topic (Lowndes et al., 2017), included in the references list. Probably, there are other papers around the topic, such as that abovementioned of Borja et al. (2019) and few others (e.g. van Oudenhoven, A. P. E., M. Schröter, R. de Groot, 2016. Open access to science on ecosystem services and biodiversity. International Journal of Biodiversity Science, Ecosystem Services & Management: 1-3).

*Thank you for the suggestions. Lowndes et al. 2017 was a formative paper that heavily influenced our thinking about this topic and we are happy to see it acknowledged. We have also added a citation to van Oudenhoven et al. 2016 in our response to the previous comment.*

Experimental design

The article content is within the Aims and Scope of the journal. Regarding the methods described, I think that they need more explanations:  
- Line 87: for the whole section, I cannot see later in the manuscript the results obtained from that search in Google Scholar. Why including this section, if there is no a section on the results? (maybe I have missed something).

*This section was included in accordance with the publishing guidelines of PeerJ for review papers, specifically to ensure “comprehensive and unbiased coverage of the literature”. Specifically, we followed the guidelines here* [*https://peerj.com/about/policies-and-procedures/#discipline-standards*](https://peerj.com/about/policies-and-procedures/#discipline-standards)*. Please note that there is no requirement for a similar section provided in the results, as suggested, so we have not included this content. We also do not provide results because of limited examples specific to bioassessment, as noted in the main text. Providing such a section would detract from our current format of the article that demonstrates basic open science principles and their application to bioassessment as a guide for others.*

- Lines 104-107: a justification of the selection of these terms is necessary. There are many other possible that could expand the number of potential papers. Some few examples: ‘open databases’, ‘assessment’, ‘ecological status’, ‘health status assessment’, etc.

*We agree that the listed terms we provide are not exhaustive and only specific words are provided that are relevant to the paper. We feel that inclusion of all terms in the text is inappropriate for this paper since we are not conducting a systematic analysis, but rather a general survey of the literature for a scoping review in accordance with the definitions provided in the link in the previous response.*

- Lines 108-109: I agree that they are scarce, but an example that the terms used are insufficient is the fact that the first paper abovementioned was not found.

*We agree with this statement, but please see our response to the previous two comments. Moreover, the inclusion of the suggested references provided by the reviewer in our revision has improved our coverage of relevant literature. We thank the reviewer for the suggestions and hope their inclusion is satisfactory.*

Although the review is organized logically, into coherent paragraphs/subsections, I cannot see why those subsections (what is the logic behind that organization, or what is the source to organize it in that way).

*There is no specific journal requirement for recommended sections of reviews, so we have organized the content to help facilitate our goals for the review. To clarify, we have added a sentence at the end of the paragraph that begins on line 116: “We structure the review by first introducing open science principles, then describing how these principles could be applied to bioassessment (i.e., developing goals, curating data, and applying open-source software) including a case study example, and lastly providing a discussion of limitations and opportunities to better contextualize real world applications of open science.”*

Validity of the findings

Some comments about the main text body.  
- Line 134: Figure 1: You should add the word ‘open’ in ‘publish data’ and ‘publish code’. In addition, in this conceptual workflow I think that something is wrong: after the publication in open access, you link the arrow with ‘conceptualize project’. However, if you are discussing about bioassessment, monitoring and taking decisions to improve aquatic systems, I think that, after publication, at least two boxes should be added: one on taking management measures, and another on recovery of the systems monitoring, which can lead to new ‘conceptualization of projects’. Otherwise, this figure is not useful, and endogamic for the research community.

*We have modified Figure 1 with the appropriate suggestions. We agree that the previous version was too generic and not specific to the bioassessment community.*

- Lines 167-170 and 188-196: an example of this are the metabarcoded databases of species (e.g. BOLD, GenBank,…).

*The following was added to line 170: “Prominent examples that can benefit next-generation bioassessment methods, such as molecular-based techniques for species identification, include the BarCode of Life Data Systems (BOLD) and GenBank repositories.”*

- Lines 253-255: there are also many taxonomical resources that can be mentioned e.g. WoRMs

*Line 255 was revised: “Species names also change regularly requiring updates to standard taxonomic effort (STE) tables that are critical for many biological indices, although some standardized databases have facilitated broad-scale comparisons (e.g., World Register of Marine Species; Costello et al. 2013).”*

- Lines 400-417: you can mention also that in Europe any project supported by public funds must provide open access data and publish all papers in open access: <https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm>. The same in Canada: <https://www.ic.gc.ca/eic/site/063.nsf/eng/h_F6765465.html>, and probably in many other countries/regions. Avoid being so local.

*The following was added to line 558: “Internationally, institutions in Europe and Canada that have projects supported by public funds are obligated to publish data and papers as open access (Horizon 2020, Tri-Agency Open Access Policy).”*

- Line 424: spell out, in the legend of Figure 3, the abbreviation: TDML

*TMDL acronym spelled out in figure caption.*

The conclusions are coherent

Comments for the Author

The manuscript has some interest, but I have identified some weaknesses, regarding the lack of connection between the methodology presented and the absence of results associated to the search in Google Scholar. In addition, some references that I have included above could be of interest. The authors should present the applicability more global, not as local as California (although it can serve as example, I have given more examples in other areas)

*Thank you for your detailed comments. Please see our responses above that address your concerns.*

**Reviewer 3:**

Comments for the Author

This is an important and innovative paper that highlight the issue of the accessibility of scientific results and data to facilitate the bioassessment of environmental impacts. The paper is very timely as the open science is becoming a standard for most of biological studies and it’s important that this practice is also adopted by ecological studies of practical importance for environmental managers.

*Thank you, we appreciate your comments and have carefully considered your suggestions below.*

The review is very well written, and I fully recommend its publication. I would suggest however few points to be considered in the revision of this paper.  
  
The first point concerns the new ecogenomic tools that are currently transforming the field of biomonitoring and bioassessment. I was quite surprised to find no reference to these tools despite the fact that their impact on so-called “democratization” of bioassessment practices is well documented (e.g. Deiner et al. 2017, Pawlowski et al. 2018, Hering et al. 2019). In particular, it is worth to be mentioned that the access to genomic or metagenomic data can be much easier and more open compared to the classical morphotaxonomic data, which are more depending on personal expertise and more difficult to compare with others. The authors could address this point when they are discussing the issue of species identification (subchapter Curating bioassessment data).

*This is an important suggestion as we avoided discussing these techniques in the original draft to focus more on conventional bioassessment methods that rely on morpho-taxonomic identification. However, we agree it would be useful to include some discussion about molecular approaches and the connection to open science. We have added some content to line 257:*

*“In contrast, molecular techniques based on DNA barcoding eliminate the need for direct species collection and morphological identification (Deiner et al. 2017; Hering et al. 2018). These next-generation approaches have capitalized on advances in database development that allow open access by diverse researchers across disciplines and are well-suited for the development of additional open science tools. Despite these advances, molecular-based approaches have also suffered from challenges related to standardization of workflows and coverage of reference databases (White et al. 2014; Elbrecht et al. 2017).”*

*Please note that we have also addressed some of these concerns in our response to reviewer 2 (additions to lines 170 and 255).*

*This paragraph was also added to the conclusions section.*

*“Efforts to formally recognize and integrate open science in bioassessment are needed now more than ever. The transition of bioassessment from taxonomic-based indices to molecular approaches presents novel challenges that will only increase in severity as researchers continue to refine methods for molecular applications (Baird and Hajibabei 2012). Although molecular-based indices share similar assessment objectives as conventional indices, the data requirements and taxonomic resolution are substantially more complex. Bioassessment researchers developing molecular methods are and will continue to be inundated with data from high-throughput DNA sequencers. Systematic approaches to document, catalog, and share this information will be required to advance and standardize the science. Molecular approaches are also dependent on existing reference libraries for matching DNA samples for taxonomic identification. The integrity of reference libraries depends greatly on the quality of metadata and documentation for contributed samples. Open science principles should be leveraged in this emerging arena to ensure that new bioassessment methods continue to have relevance for determining the condition of aquatic resources.”*

The second point, somehow related to the first, concerns a certain conservationism of bioassessment practitioners regarding scientific progress. As discussed by Guareschi et al. 2019, this is not only about new tools but also about ecological changes, such as the presence of alien species, or changes in taxonomic identification. Of course, open science can help pointing out these changes and interpreting them correctly. However, even the most accessible research will do nothing if the environmental managers are not particularly interested in getting updated about the new research advances and implementing them in their practice.

*This is an insightful comment that speaks to both the challenges of developing relevant assessment methods and the need to encourage their timely application in management scenarios. We agree that open science can indeed facilitate these advances by improving workflows that allow for reproducibility and rapid update of methods (e.g., in the case of taxonomic changes or updates to species lists with invasions), while also allowing improved ability to translate the science in an applied setting. The latter will always be a challenge and one of our goals for developing this paper is to address this need by encouraging researchers to adopt open science principles into their workflows (e.g., lines 90, 178, 474, 482-488). We also specifically emphasize the role of administrative leaders (e.g., managers) in encouraging the adoption of open science to address this concern (lines 503-513). We feel these sections address the challenge of science translation and adoption in applied settings through use of open science.*

My final critical remark concerns the scope of the review and its emphasis on open source software applications. The authors provide a very exhaustive description of using R for bioassessment, but they go into a lot of details that somehow contrast with a general character of the review. They assume most of environmental managers know how to use R, which I doubt is true. In fact, the authors seem to agree with me, when they point out that the education and training are the main limitations of the access to the open science tools and data. It’s probably the main message of this review paper and I would rather focus on it than on describing the various applications of tools that most probably remain inaccessible for most of practitioners.

*This section is given extra emphasis because it highlights R as a critical resource that can be leveraged to embrace and facilitate open science principles. Our intent was to demonstrate its applicability, while also providing an inventory of relevant packages within the R environment that bioassessment practitioners can use. The current section is approximately 1132 words. We have revised the R content to reduce the number of words to 1016. Further, our existing text in the “Limitations and opportunities” section addresses several of the concerns in the reviewer’s comments, i.e., training is necessary, encourage a community of practice, etc.*

References:  
Hering et al. 2019  
<https://www.sciencedirect.com/science/article/pii/S0043135418301830?via%3Dihub>  
Guareschi et al. 2019  
<https://www.sciencedirect.com/science/article/pii/S0048969719300087?via%3Dihub>  
Deiner et al. 2017  
<https://onlinelibrary.wiley.com/doi/full/10.1111/mec.14350>