

**An approach to establishing a workflow pipeline for synergistic analysis of osteological and biochemical data. The case study of Amvrakia in the context of Corinthian colonisation between 625-189 BC in Epirus, Greece.**

**Xanthopoulos Kiriakos, Georgiadou Aggeliki, Papageorgopoulou Christina**

### **Summary**

Overall, this is a welcome contribution to the field of bioarchaeology, both in regards to material from the Aegean and beyond. The study's technical merit and scientific significance is unquestionable. The field could greatly benefit from such a remarkable idea and I will be looking forward to the final publication. However, I would recommend some minor revisions to the manuscript in order for it to be more precise and complete. Thus, please accept some suggestions for improvements that I hope would greatly improve the paper.

### **Title and abstract**

In more detail, the title and the abstract suggest an innovative and promising idea that is greatly lacking in the field of bioarchaeology. They both provide an adequate synopsis of the manuscript as submitted and a comprehensive outline of the content and the aims of the paper. No results are provided, but this does not pose a problem for the present study and its scope.

### **Introduction**

The introduction refers to the limitations the current state of research faces regarding data collection and standardization. However, it does not contain sufficient literature review on the matter and suggest any publications that have been trying to overcome such limitations.

Following there are some publications towards this direction:

Standardized Osteological Database (SOD), Annual Review Anthropology 1996 article

<https://www.jstor.org/stable/2155819?seq=17>

Wellcome Osteological Research Database

<https://www.museumoflondon.org.uk/collections/other-collection-databases-and-libraries/centre-human-bioarchaeology/about-osteological-database>

Caruso, A., Karligiotti, A., Selempa, G. and Nikita, E., STARC OSTEOARCH: An open access resource for recording and sharing human osteoarchaeological data. International Journal of Osteoarchaeology. <https://airtable.com/shr4mDZga3uMFN35n>

### **Paleopathology Association efforts for standardization**

Rose JC, Anton S, Aufderheide A, Buikstra JE, Eisenberg L, et al. 1991. Paleopathology Association Skeletal Database Committee Recommendations. Detroit: Paleopathology Association.

White, W. 2008. Databases. In *Advances in Human Paleopathology*, ed. R Pinhasi and S. Mays, 177-188. John Wiley and Sons. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470724187.ch9>

### **Methods, Analyses and Publication**

The procedures regarding decision-making and the approach implemented are stated clear enough. However, the methodologies implemented need to be clarified in more precision. Furthermore, even though the authors are using the colonization of Ambrakia as a case study, they do not offer specific examples coming from the application of their methodologies in the case study. E.g. the way they correlate oral pathologies with isotopic values, as they proposed on line 59 of the introduction. Therefore, the suggested methodology for a pipeline of workflow analysis based on an Ambrakia, as the abstract and title suggest, is merely given at a theoretical level, while the practical part is critically missing from the study. Such an addition would greatly help the scientific community understand and implement the authors' methods in order to facilitate data sharing and correlation.

The sections concerning Recording Data, Importing and Tidying Datasets, Analyses and Publication are not very clear and could greatly benefit from some additional details.

Thus, I would like to suggest the following additions:

- a) The names of the parameters the authors utilize and the analyses they are related to should be given in detail.
- b) The specific descriptive statistics and/or specific statistical tests that are utilized should be mentioned.
- c) The specific R packages the authors utilize should be mention in more precision, e.g. the authors refer to tidying datasets. Is this a reference to the TidyR R package?
- d) In the beginning of the paper the authors mention R Scripts. For the task the authors are describing R markdown notebooks seem appropriate. Have the authors considered them?
- e) Figure 3 could be replaced with a more detailed workflow diagram.

Finally, the authors provide no discussion regarding tools for data querying tables (such as SQL) in order to discover correlations in relation to age, sex, archaeological data and osteological biomarkers that are necessary for researchers to identify, interpret, and report. Even though this does not suggest a mandatory revision for the present study, its consideration could potentially benefit the paper.