Review of: "AI-Generated Hallmarks of Aging and Cancer: A Computational Approach Using Causal Emergence and Dependency Networks"

Susan Mertins[1,2]

1 Frederick National Laboratory for Cancer Research, Frederick, United States
2 BioSystems Strategies, LLC, Frederick, United States

Potential competing interests: No potential competing interests to declare.

The published report on the utilization of meta-type hallmarks for aging and disease is an interesting concept that may indicate a better understanding of biological systems and any disease within them.

While efforts should be praised for the attempt to utilize existing large data sets, many biological questions could be considered in the original premise and applications to the results.

In the original hypothesis, it would be helpful to indicate why quantifying DNA methylation as a marker of aging is valid. For example, there will be other marks on DNA that counter methylation, and the complexity of the transcription machinery should be considered and/or described. Citing concrete examples from the literature would be a good first start. Further, what other data sets reported in the literature might be useful for defining aging?

Regarding the methodology, it is of concern that healthy controls (as described) are from a young population and do not describe any given disease data set. While it is clear that a young healthy control population is needed to define aging hallmarks alone, overall, an age-matched healthy control population is needed for cancer, depression, and COVID-19 severity. These are morbidities not necessarily defined by age at diagnosis.

And relatedly, it would increase the understanding behind the rationale for this study to know the percent of the aged populations that have the co-morbidities that appear to be expected by the author. For instance, what percent of the aged population suffers from Alzheimer’s disease and cancer?

Other biological considerations would center on the applications the author suggests may become useful given their algorithms. For example, does the author expect that in the clinic, a blood sample will be drawn, then be utilized to identify a drug to treat the co-morbidity in a more effective fashion? A specific example within the study would demonstrate this and aid the reader. Perhaps there already are existing pharmaceuticals that address this.

In light of FAIR principles, are the code and data sets available to the scientific community?

Finally, overall, the completely computational nature of the study is a clear limitation, and future studies should be suggested. The tyranny of numbers in such a large analysis raises concerns for correlations that will be identified but not biologically meaningful.

In sum, the report here is of interest from both a basic and applied aspect, and deeper details as follow-up will be useful to the scientific community.