doi 10.18699/bgrs2024-13-23

Developing a standardized methodology to validate epigenetic aging clocks

Efimov E.[1*†], Kriukov D.[1, 2†], Kuzmina E.[1, 2], Dylov D.V.[1, 2], Khrameeva E.E.[1]

[1] *Skolkovo Institute of Science and Technology, Moscow, Russia*

[2] *Artificial Intelligence Research Institute, Moscow, Russia*

** Evgeniy.Efimov@skoltech.ru*

† *These authors contributed equally to this work*

Key words: aging; biomarkers; epigenetics; aging clocks; benchmarking; machine learning

Motivation and Aim: Developing a measure of human aging process becomes more and more valuable, as emerging longevity therapies require this measurement to be feasibly evaluated in clinical trials [1, 2]. However, contrary to the common machine learning models, predictors of biological age, so-called aging clocks, cannot be compared between each other in terms of their correlation with chronological age, since high accuracy of chronological age prediction does not imply high accuracy at estimating biological age. In our work, we aimed to develop a methodology to validate aging clocks by testing their ability to identify accelerated aging in a standardized panel of aging-accelerating conditions (AACs).

Methods and Algorithms: Relying on the published studies of age-related conditions, we elaborated an evidence-based set of putative AACs. To showcase our methodology, we collected a large number of open-access, age-annotated, and microarray-based datasets of human blood DNA methylation using the NCBI Gene Expression Omnibus database. We also extracted published epigenetic clock models from the respective studies and included them into our comparison pipeline. To benchmark the clock models, we proposed four tasks: two-sample aging acceleration prediction (where both healthy and AAC cohorts are available), one-sample aging acceleration prediction (in datasets containing the AAC cohort only), chronological age prediction accuracy, and systematic chronological age prediction bias. To summarize all tasks, we also suggested a cumulative benchmarking score that would penalize the results of absolute aging acceleration prediction by the ratio of prediction bias to prediction accuracy.

Results: Dataset search yielded 66 individual datasets in total comprising 19 of determined AACs. We also collected and implemented 13 existing epigenetic clock models, containing both the first generation (trained to predict chronological age only) and the second generation (trained to predict time to death) of aging clocks. The two-sample task revealed that the second-generation aging clocks (PhenoAge [3, 4] and GrimAgeV1 [5, 6] variants) appear on top, particularly at predicting aging acceleration for the immune system-related diseases (such as HIV infection and inflammatory bowel disease). Some disease classes (namely, cardiovascular and metabolic diseases) were misestimated in regards to healthy controls by all tested models. In contrast to that, first-generation aging clocks by Zhang et al. [7] and Hannum et al. [8] emerged the best at the one-sample task, in addition to the GrimAge variants [5, 6], across multiple disease classes. However, one-sample score might be compromised by systematic model error. Therefore, it was not surprising to find that, according to the cumulative benchmarking score, second-generation aging clocks still outperformed other models.

Conclusion: In our work, we developed the first systematic pipeline for validating epigenetic aging clocks based on public data to foster more rapid and robust advancements in the field of biological age research.

Funding: This work was supported by Program "Skolkovo Institute of Science and Technology – University of Sharjah Joint Projects: Artificial Intelligence for Life".

References

1. Moqri M. et al. Validation of biomarkers of aging. *Nat Med.* 2024;30(2):360-372. doi 10.1038/s41591-023-02784-9
2. Schork N.J. et al. Does Modulation of an Epigenetic Clock Define a Geroprotector? *Adv Geriatr Med Res.* 2022;4(1):e220002. doi 10.20900/agmr20220002
3. Levine M.E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY).* 2018;10(4):573-591. doi 10.18632/aging.101414
4. Higgins-Chen A.T. et al. A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. *Nat Aging*. 2022;2(7):644-661. doi 10.1038/s43587-022-00248-2
5. Lu A.T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY).* 2019;11(2):303-327. doi 10.18632/aging.101684
6. Lu A.T. et al. DNA methylation GrimAge version 2. *Aging (Albany NY).* 2022;14(23):9484-9549. doi 10.18632/aging.204434
7. Zhang Q. et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Med.* 2019;11(1):54. doi 10.1186/s13073-019-0667-1
8. Hannum G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359-367. doi 10.1016/j.molcel.2012.10.016