More than just a number: modelling biological aging and vulnerability

Sluiskes, M.H.

Chapter 1

Introduction

1.1 Biological age 101

Many of us want to grow old, few want to *be* old. Being old implies having aged, and the aging process is a phenomenon that most of us do not particularly like: it entails a gradual functional decline, the onset of age-related chronic diseases and eventually, death [1]. As each individual is aging, so is the world as a whole: the global population is growing old at an unprecedented rate. Current predictions are that by the year 2100, the number of people aged 65 and up will have exceeded the number of people aged under 20 by about half a billion [2]. An aging global population implies an increasing percentage of those suffering from age-related health conditions, which in turn puts a large burden on healthcare systems.

Given the challenges that come with aging, both on a personal as well as on a societal level, obtaining a better insight into aging is highly relevant. Developing models that can accurately and reliably predict how far someone has progressed in the aging process is an important part of this. From an inference point of view, such aging status predictors can help to elucidate underlying drivers of aging. From a prediction point of view, they may contribute to identifying individuals that are most at risk for developing age-related diseases, to make sure they get adequate and timely care.

The most popular and straightforward approach to obtain an idea of one's aging status is to consider their chronological age, which is simply defined as time alive since birth, generally given in number of full years. Chronological age, also known as calendar age, is a major risk factor for functional impairments, chronic diseases and mortality [3]. However, it is far from a perfect aging status predictor: individuals of the same chronological age show considerable variation in their health outcomes [4, 5].

Given the heterogeneity in aging outcomes of people of the same chronological age, the idea has risen that individuals age at different rates, and that this variability in aging rate is inadequately captured by chronological age. As a result, it has been suggested that individuals possess a *biological age* in addition to their chronological age. This biological age reflects someone's true aging status and is hence a better predictor of aging-related outcomes (such as the onset of age-related diseases and death) than chronological age. The difference between biological and chronological age, also known as the 'age gap', is often denoted by the symbol Δ. If biological age exceeds chronological age, the age gap Δ is positive, which is indicative of accelerated aging. Individuals for which this holds are expected to have a lower remaining lifespan and

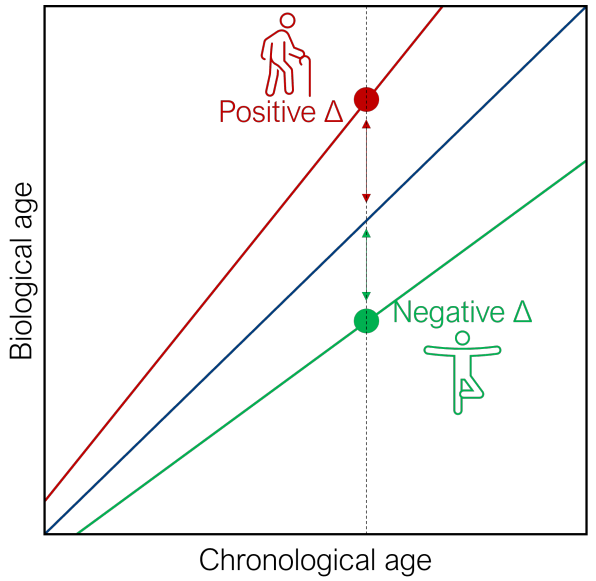

Figure 1.1: Illustration of biological age. If biological age exceeds chronological age, this is indicative of accelerated aging. The age gap Δ will be positive. If biological age is lower than chronological age, this is indicative of slower aging. The age gap Δ will be negative.

are at an increased risk of developing age-related diseases. In contrast, if biological age is lower than chronological age, Δ is negative and individuals for which this is the case are considered to be aging more slowly. Figure 1.1 contains a visual representation of the relation between chronological age, biological age and the age gap Δ.

Biological age is a latent concept: it cannot be directly measured. This complicates the construction and evaluation of predictors of biological age. Without a clear definition of what biological age is, different biological age predictors might capture different aspects. This makes it difficult to compare predictors and decide which ones are performing best. The concept of biological age should therefore be operationalized a priori. In addition, the statistical methods that are subsequently used should be carefully chosen and their underlying assumptions well understood, to ensure that they are in line with how the concept of biological age is regarded within the field of aging. If currently existing statistical methods do not suffice, new ones have to be developed.

These methodological challenges illustrate that statisticians have an important role to play within the field of biological age prediction, making this topic inherently

multi- and interdisciplinary. The accuracy, interpretability and reliability of biological age predictors strongly hinges on the underlying statistical methods. This thesis therefore approaches aging from a statistical perspective. It starts by discussing the assumptions behind commonly used biological age prediction approaches and pointing out their flaws. In addition, it introduces two new prediction approaches: a framework for biological age prediction based on a proper operationalization of biological age and a statistical model for predicting age-related multi-morbidity.

1.2 A brief history

The limitations of chronological age as a measure for one's underlying overall aging status have long been acknowledged. One of the first to discuss aging beyond just the passage of time, introducing the idea of biological age, was Harry Benjamin [6]. To obtain a better measure of true aging status, he proposed a so-called "health inventory for biologic age estimation". This inventory had to be completed by a physician, through the examination of various health factors of an individual. A certain deviation from chronological age had to be manually assigned by the physician, based on the perceived health status. As Benjamin himself admitted, "The importance of the various factors is by necessity rather arbitrarily appraised". Nevertheless, his view on the concept of biological age was in many ways close to how the subject is approached today: "It is the true age determined by heredity, state of health, and functioning capacity. It can be higher one year and lower the next, if health has improved".

The next step in biological age prediction was taken almost two decades after Benjamin. Hollingsworth et al. [7] presented a predictor of biological age based on nine different physiological functions that change with chronological age (such as skin elasticity or blood pressure). Data on approximately 450 Japanese Hiroshima survivors was used, as their aim was to check for evidence of accelerated aging among these survivors. A standard linear regression model was fitted for each of the nine predictor variables, using chronological age as the outcome of interest, after which the nine model predictions were combined into one overarching biological age prediction. This was one of the first attempts of using a statistical model to predict biological age and detecting aging gap differences between individuals of the same chronological age.

This line of thought was continued by Alex Comfort, who is often considered to be one of the founding fathers of the field of biological age, laying the methodologi-

cal and scientific groundwork for the many predictors that were to follow [8]. In his seminal paper on measuring the aging rate in men, he proposed a test battery (i.e. a set of tests to obtain an assessment of a particular phenomenon) consisting of a wide variety of potential predictors of biological age [9]. He also provided a detailed description of the test procedure. Comfort was one of the first to refer to biological age predictors as 'age clocks', which has proven to be a strong and persistent metaphor: the terms '(biological) age predictors' and 'aging clocks' have since then been used practically interchangeably.

In the decades thereafter, more attempts were made to establish a set of variables that could together predict biological age [10, 11, 12]. Such variables were called '(bio)markers of aging'. Biological age was calculated based on these (candidate) markers of aging, by measuring the markers in an individual and comparing them to age norms. These age norms were determined by conducting a multiple linear regression with chronological age as the outcome, using the candidate markers as predictors. Biological age was then defined as predicted chronological age.

In the 21st century, thanks to the advent of high throughput bio-molecular technologies, attention started to shift from low-dimensional clinical variables to high-dimensional omics-based variables as potential biomarkers of aging. DNA methylation, which refers to the addition or removal of methyl groups to the DNA, thereby causing changes that affect the way genes work, proved to be a particularly promising predictor of chronological age. Two famous DNA methylation-based age predictors are the clocks by Hannum et al. [13] and Horvath [14]. The statistical methodology used to build these clocks remained unchanged: chronological age was taken as the outcome of interest, after which a (in this case penalized) multiple linear regression model was fitted using the DNA methylation data as the predictor variables. The difference between predicted chronological age and true chronological age was taken to be indicative of one's biological age.

The publication of Hannum's and Horvath's epigenetic clocks marked the beginning of a new era in biological age prediction. Many more biological age predictors were built, using different types of omics predictor variables, such as metabolomics [15], proteomics [16], transcriptomics [17], glycomics [18] and microbiomics [19]. These clocks were all constructed using the same statistical approach of regressing chronological age on the candidate markers of interest.

Validation of biological age clocks is not straightforward, as biological age cannot be directly observed. Therefore, the common approach within the biological age field is to evaluate whether the difference between predicted chronological age (=

predicted biological age) and true chronological age, denoted by $\hat{\Delta}$, is associated with time-to-mortality and other age-related outcomes. The premise is that those with a positive predicted age gap $\hat{\Delta}$ (i.e. where predicted biological age exceeds true chronological age) have aged more than what might be expected based on their chronological age alone.

In the last five years, the most important development have not only been biological—the popularity of omics-based predictors is still steadily growing—but methodological as well. Instead of fitting regression models with chronological age as the outcome ('age-trained'), new models have been proposed that use time-to-mortality as the outcome ('mortality-trained'). This requires longitudinal data in which subjects are followed over time, whereas cross-sectional data sufficed for the age-trained clocks. Biological age predictors based on time-to-mortality are known as 'second generation' predictors [20]. The most well-known second generation predictors are GrimAge [21, 22] and PhenoAge [23]. Both these mortality-trained predictors use DNA methylation data as predictor variables in the model. They were found to outperform the first generation age-trained clocks: their associated $\hat{\Delta}$ is more strongly and persistently associated with various health conditions, physical and cognitive performance, age-related clinical phenotypes and frailty measures [24, 25, 20, 26].

The third generation of biological age clocks does not yet exist, but has been speculated to involve repeated measures of candidate markers of interest [20]. This means that changes in these markers can also be incorporated in the model, instead of only considering a single value at a single time point. Research in this direction is still in its early stages, which can probably partly be attributed to the fact that datasets with many repeated measurements of relevant omics-predictors are still relatively scarce. Instead, many of the recent developments seem to have occurred on the side of the model outcome: instead of only considering time-to-mortality, some recent work focused on (also) considering time to various age-related diseases [27, 28]. Datasets with multiple age-related outcomes are more abundant than rich repeated measures datasets, thanks to the advent of large-scale long-term population cohort studies such as the UK Biobank. Such studies often also use data from electronic health records and hence contain information on various age-related events. However, most multi-outcome models are either focusing on inference or do not (yet) predict on an age-scale and are therefore strictly speaking not biological age predictors.

1.3 Known methodological issues

Both first generation clocks, using cross-sectional data, and second generation clocks, using time-to-event data, have been criticized for several methodological reasons. This section briefly summarizes the most important concerns.

Most criticism pertains the first generation clocks. More than 25 years ago, Hochschild [29] pointed out several statistical flaws of these chronological age-trained predictors. Firstly, he discusses the so-called "biomarker paradox": included candidate biomarkers are weighted according to their strength of correlation with chronological age, which is the regression's outcome variable. Any hypothetical biomarker that perfectly predicts chronological age would be useless in predicting biological age, as first generation methods rely on the differences between true chronological age and predicted chronological age to be informative of an individual's rate of aging. It is unclear how well any model should predict chronological age in order to be informative of biological age: it is not a matter of the more precise, the better.

Additionally, a biomarker that is associated with chronological age may not be predictive of individual aging. This argument goes two ways: candidate markers may be associated with chronological age but may not provide any information on one's age gap Δ. An example of this could be the degree of baldness in men: this is undoubtedly associated with chronological age, but men who grow bald earlier in life do not show other signs of accelerated aging, nor do they die sooner. Conversely, candidate markers may not be associated with chronological age but may still be informative of the individual age gap Δ. A simple regression of chronological age on a set of candidate markers of aging will miss such markers.

Another well-known problem, also discussed by Hochschild, is the regression to the mean phenomenon. In all regression models, predicted values tend towards the mean of the outcome variable in the data set on which the model was fitted. This occurs whenever outcome and predictor(s) are not perfectly correlated. Hence, due to this phenomenon true low outcome values will be overestimated and true high values will be underestimated. As the difference between predicted chronological age and true chronological age, the predicted age gap $\hat{\Delta}$, is often taken to be indicative of the difference between true biological age and chronological age, this means that people younger than the training data sample will on average have a positive $\hat{\Delta}$, and those older than the training data sample will on average have a negative $\hat{\Delta}$. This pattern does not reflect a biological phenomenon but is purely a statistical artefact.

A final point of critique on the first-generation clocks concerns the way they are

(over)interpreted. Many candidate biomarkers of interest tend to be strongly correlated. This might result in the phenomenon that for a given biomarker, its coefficient sign in a univariate regression has the opposite direction of its sign in a multiple regression. Zooming out, this can be seen as an implicit warning against a causal interpretation of model coefficients. The (impossibility of a) causal interpretation of biological aging clocks remains an active field of research and discussion today [30, 31].

Notwithstanding these fundamental points of criticism, cross-sectional clocks have remained popular today [32, 33, 34, 35]. This can probably at least partly be attributed to the fact that cross-sectional data is much easier to obtain than longitudinal data. However, it is to be expected that the shift away from first generation age-trained clocks towards second generation mortality-trained clocks will continue in the years to come. These second-generation clocks overcome many of the issues inherent to first-generation clocks (primarily those caused by using chronological age as the regression outcome), although other concerns remain. Most importantly, the two most popular second-generation clocks, GrimAge and PhenoAge, are not based on a formal operationalization of biological age. Their prediction approaches hence remain fairly ad hoc and a proper statistical estimand is missing. This makes it difficult to properly evaluate whether predictors are capturing what they (cl)aim to capture and to compare and contrast different methods. Another limitation of the second-generation clocks is that they are trained on time-to-mortality only. This can be considered a too restricted view on what it means to age: lifespan is not all that matters, healthspan is just as important.

1.4 Disentangling excess mortality risks in vulnerable groups

Although this thesis primarily discusses methodological developments for biological age prediction, one chapter has a different focus: it introduces a new methodological framework for estimating excess mortality (e.g. a higher mortality risk than expected) among vulnerable groups during a period of crisis.

While distinct from biological age estimation, the topics are connected: biological age aims to capture vulnerability on the individual level, whereas excess mortality aims to capture vulnerability on the population level. In addition, the excess mortality chapter is linked to the biological age chapters through its real data application,

in which we focus on excess mortality among elderly care home residents. This is closely connected to the study of vulnerability in aging.

To estimate excess mortality, a popular measure is *relative survival*. This captures how much the specific condition or circumstance of the vulnerable group of interest affects survival relative to the general population. Relative survival models assume that the total observed mortality hazard (the instantaneous probability of death) at time t, $\lambda_{tot}(t)$, is the sum of some background hazard, $\lambda_{background}(t)$ plus the excess hazard $\lambda_{excess}(t)$. The background hazard is usually taken from existing population lifetables, stratified by age, sex and birth year.

Relative survival models are typically used to model the excess risk of a specific vulnerable subgroup of the general population: for example, to compare the overall survival of breast cancer patients with that of the general population in a given period of time. Hence, the observed mortality of two *different* populations (breast cancer patients and non-breast cancer patients) is compared in the *same* period of time. In addition, relative survival models can be used to model the excess hazard of the *same* population in *different* time periods. This can for instance be useful when the entire population experiences a period of crisis resulting in excess mortality, such as a war or pandemic. The survival of the general population before and during the crisis period can be compared to obtain an estimate of the excess mortality experienced by the general population during the crisis.

A recent example of such a crisis is the Covid-19 pandemic, where excess mortality was observed on the level of the entire population [36, 37]. It is reasonable to expect that vulnerable subgroups suffered more during the Covid-19 pandemic than the general population. A particularly vulnerable subgroup during the Covid-19 pandemic were care home residents [38, 39]. To accurately model the Covid-19-related excess mortality that can be attributed to being a care home resident, both the pre-existing vulnerability of this group and the Covid-19-related excess mortality experienced by the entire population should be accounted for.

To this end, this thesis introduces the novel concept of "excess excess mortality" and presents a framework to estimate it. This concept is useful in scenarios where there is more than one cause or indicator of excess mortality, where their might be an interaction between these two causes, resulting in more or less excess mortality than when considering the excess mortality sum of both individual causes. In this thesis, we define this additional excess mortality resulting from the interaction as the "excess excess" mortality. Concretely, this thesis illustrates how the concept of excess excess mortality is useful when disentangling the different excess mortality components of

Dutch care home residents during the Covid-19 pandemic.

1.5 Main contributions and outline of this thesis

This thesis has two aims. Its main aim is to advance the understanding and prediction of (biological) aging, by incorporating statistical rigor and introducing innovative methodologies to address current limitations in the field. Its second aim is to extend existing relative survival methodology in order to properly estimate the different excess mortality components experienced by vulnerable subgroups of the general population during periods of crisis.

The assumptions behind cross-sectional first-generation biological age predictors are discussed and critically assessed in **Chapter 2**. We show that these predictors all rely on the same strong yet untestable assumption, namely namely that a candidate marker of aging's association with chronological age is directly informative of its association with the age gap Δ. If this assumption does not hold, weights assigned to candidate markers of aging are uninformative. Synthetic and real data are used to illustrate the assumption and its consequences if it does not hold. Although limitations of cross-sectional biological age prediction methods have been acknowledged before, the issues with the underlying assumption have never been explicitly discussed. This chapter serves as a warning against the use of cross-sectional clocks and encourages caution in their use.

The link between biological age prediction and the regression to the mean phenomenon, and how ignoring this phenomenon can lead to premature or even invalid conclusions, is discussed in **Chapter 3**. This concise chapter explains regression to the mean in simple terms and uses a synthetic data example to illustrate which biased conclusions may arise when directly comparing the difference between true and predicted chronological age $\hat{\Delta}$, often taken to be indicative of accelerated or decelerated biological aging, between groups with different ages.

Chapter 4 presents *AccelerAge*: a new methodological framework to predict biological age with time-to-event data using Accelerated Failure Time (AFT) models. This is hence a second generation predictor, but one that is based on a formal operationalization of the concept biological age, as opposed to currently existing second generation predictors. The AccelerAge framework directly models the effect of candidate predictors of aging on an individual's survival time, aligning with the prevalent metaphor of aging as a clock. In addition, underutilized evaluation measures for assessing the performance of biological age predictors are introduced. We compare

predictors based on the AccelerAge framework to a predictor based on a GrimAge-type predictor, using simulated data as well as data from the UK Biobank and the Leiden Longevity Study.

In **Chapter 5** a step towards a third generation predictor of biological age is taken. Second generation predictors use time-to-mortality as their outcome of interest, but aging is a multifaceted process that entails more than lifespan alone. A new model is hence introduced that can deal with multiple time-to-event outcomes and highly correlated or high-dimensional predictor variables, namely the penalized reduced rank regression model for multi-outcome survival data. By imposing a rank constraint on the coefficient matrix, the reduced rank regression model can identify shared latent factors driving multiple outcomes. In this thesis a lasso-penalized reduced rank regression model is fitted on UK Biobank participants, using over 200 metabolic variables as predictors and the onset of seven age-related diseases and mortality as the outcomes of interest. This results in a metabolite-based score of age-related disease susceptibility and mortality.

The final contribution of this thesis, contained in **Chapter 6**, has a slightly different scope than the rest of this thesis. Here, the focus is not on biological age, but on quantifying different components of excess mortality in vulnerable populations during crises. The concept of excess excesss mortality is introduced to capture the additional mortality risks faced by vulnerable subgroups of the general population, beyond what can be explained by the background mortality, their excess mortality due to being vulnerable and the general population's excess mortality due to the crisis. This offers a new perspective on how to model and interpret excess mortality in vulnerable groups during crises. An additive hazards model is used to disentangle the different components of (excess) mortality. Using individual-level whole population data from Statistics Netherlands, the excess excess mortality of Dutch care home residents aged 70+ during the Covid-19 pandemic is modelled.

1.6 Bibliography

[1] Carlos López-Otín, Maria A Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. Hallmarks of aging: An expanding universe. *Cell*, 186(2):243–278, 2023.

[2] Stein Emil Vollset, Emily Goren, Chun-Wei Yuan, Jackie Cao, Amanda E Smith, Thomas Hsiao, Catherine Bisignano, Gulrez S Azhar, Emma Castro, Julian Chalek, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the global burden of disease study. *The Lancet*, 396(10258):1285–1306, 2020.

[3] Brian K Kennedy, Shelley L Berger, Anne Brunet, Judith Campisi, Ana Maria Cuervo, Elissa S Epel, Claudio Franceschi, Gordon J Lithgow, Richard I Morimoto, Jeffrey E Pessin, et al. Geroscience: linking aging to chronic disease. *Cell*, 159(4):709–713, 2014.

[4] David J Lowsky, S Jay Olshansky, Jay Bhattacharya, and Dana P Goldman. Heterogeneity in healthy aging. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 69(6):640–649, 2014.

[5] Linda Partridge, Joris Deelen, and P Eline Slagboom. Facing up to the global challenges of ageing. *Nature*, 561(7721):45–56, 2018.

[6] Harry Benjamin. Biologic versus chronologic age. *Journal of Gerontology*, 2(3):217–227, 1947.

[7] James William Hollingsworth, Asaji Hashizume, and Seymour Jablon. Correlations between tests of aging in hiroshima subjects–an attempt to define "physiologic age". *The Yale Journal of Biology and Medicine*, 38(1):11, 1965.

[8] Tiago Moreira. Anticipatory measure: Alex Comfort, experimental gerontology and the measurement of senescence. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 77:101179, 2019.

[9] Alex Comfort. Test-battery to measure ageing-rate in man. *The Lancet*, 294(7635):1411–1415, 1969.

[10] Toshiyuki Furukawa, Michitoshi Inoue, Fumihiko Kajiya, Hiroshi Inada, Seiichi Takasugi, Sugao Fukui, Hiroshi Takeda, and Hiroshi Abe. Assessment of biological age by multiple regression analysis. *Journal of Gerontology*, 30(4):422–434, 1975.

[11] Ian W Webster and Alexander R Logie. A relationship between functional age and health status in female subjects. *Journal of Gerontology*, 31(5):546–550, 1976.

[12] VP Voitenko and AV Tokar. The assessment of biological age and sex differences of human aging. *Experimental Aging Research*, 9(4):239–244, 1983.

[13] Gregory Hannum, Justin Guinney, Ling Zhao, LI Zhang, Guy Hughes, SriniVas Sadda, Brandy Klotzle, Marina Bibikova, Jian-Bing Fan, Yuan Gao, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, 49(2):359–367, 2013.

[14] Steve Horvath. DNA methylation age of human tissues and cell types. *Genome Biology*, 14(10):1–20, 2013.

[15] Erik B Van Den Akker, Stella Trompet, Jurriaan JH Barkey Wolf, Marian Beekman, H Eka D Suchiman, Joris Deelen, Folkert W Asselbergs, Eric Boersma, Davy Cats, Petra M Elders, et al. Metabolic age based on the BBMRI-NL 1H-NMR metabolomics repository as biomarker of age-related disease. *Circulation: Genomic and Precision Medicine*, 13(5):541–547, 2020.

[16] Toshiko Tanaka, Angelique Biancotto, Ruin Moaddel, Ann Zenobia Moore, Marta Gonzalez-Freire, Miguel A Aon, Julián Candia, Pingbo Zhang, Foo Cheung, Giovanna Fantoni, et al. Plasma proteomic signature of age in healthy humans. *Aging Cell*, 17(5):e12799, 2018.

[17] Marjolein J Peters, Roby Joehanes, Luke C Pilling, Claudia Schurmann, Karen N Conneely, Joseph Powell, Eva Reinmaa, George L Sutphin, Alexandra Zhernakova, Katharina Schramm, et al. The

transcriptional landscape of age in human peripheral blood. *Nature Communications*, 6(1):1–14, 2015.

[18] Jasminka Kristic, Frano Vuckovic, Cristina Menni, Lucija Klaric, Toma Keser, Ivona Beceheli, Maja Pucic-Bakovic, Mislav Novokmet, Massimo Mangino, Kujtim Thaqi, et al. Glycans are a novel biomarker of chronological and biological ages. *The Journals of Gerontology: Series A*, 69(7):779–789, 2014.

[19] Fedor Galkin, Polina Mamoshina, Alex Aliper, Evgeny Putin, Vladimir Moskalev, Vadim N Gladyshev, and Alex Zhavoronkov. Human gut microbiome aging clock based on taxonomic profiling and deep learning. *iScience*, 23(6):101199, 2020.

[20] Cathal McCrory, Giovanni Fiorito, Belinda Hernandez, Silvia Polidoro, Aisling M O'Halloran, Ann Hever, Cliona Ni Cheallaigh, Ake T Lu, Steve Horvath, Paolo Vineis, et al. GrimAge outperforms other epigenetic clocks in the prediction of age-related clinical phenotypes and all-cause mortality. *The Journals of Gerontology: Series A*, 76(5):741–749, 2021.

[21] Ake T Lu, Austin Quach, James G Wilson, Alex P Reiner, Abraham Aviv, Kenneth Raj, Lifang Hou, Andrea A Baccarelli, Yun Li, James D Stewart, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)*, 11(2):303, 2019.

[22] Ake T Lu, Alexandra M Binder, Joshua Zhang, Qi Yan, Alex P Reiner, Simon R Cox, Janie Corley, Sarah E Harris, Pei-Lun Kuo, Ann Z Moore, et al. DNA methylation GrimAge version 2. *Aging (Albany NY)*, 14(23):9484, 2022.

[23] Morgan E Levine, Ake T Lu, Austin Quach, Brian H Chen, Themistocles L Assimes, Stefania Bandinelli, Lifang Hou, Andrea A Baccarelli, James D Stewart, Yun Li, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*, 10(4):573, 2018.

[24] Robert F Hillary, Anna J Stevenson, Daniel L McCartney, Archie Campbell, Rosie M Walker, David M Howard, Craig W Ritchie, Steve Horvath, Caroline Hayward, Andrew M McIntosh, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. *Clinical Epigenetics*, 12(1):1–12, 2020.

[25] Jane Maddock, Juan Castillo-Fernandez, Andrew Wong, Rachel Cooper, Marcus Richards, Ken K Ong, George B Ploubidis, Alissa Goodman, Diana Kuh, Jordana T Bell, et al. DNA methylation age and physical and cognitive aging. *The Journals of Gerontology: Series A*, 75(3):504–511, 2020.

[26] Lieke M Kuiper, Harmke A Polinder-Bos, Daniele Bizzarri, Dina Vojinovic, Costanza L Vallerga, Marian Beekman, Martijn ET Dollé, Mohsen Ghanbari, Trudy Voortman, Marcel JT Reinders, et al. Epigenetic and metabolomic biomarkers for biological age: A comparative analysis of mortality and frailty risk. *The Journals of Gerontology: Series A*, 78(10):1753–1762, 2023.

[27] Thore Buergel, Jakob Steinfeldt, Greg Ruyoga, Maik Pietzner, Daniele Bizzarri, Dina Vojinovic, Julius Upmeier zu Belzen, Lukas Loock, Paul Kittner, Lara Christmann, et al. Metabolomic profiles predict individual multidisease outcomes. *Nature Medicine*, 28(11):2309–2320, 2022.

[28] Maik Pietzner, Isobel D Stewart, Johannes Raffler, Kay-Tee Khaw, Gregory A Michelotti, Gabi Kastenmüller, Nicholas J Wareham, and Claudia Langenberg. Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. *Nature Medicine*, 27(3):471–479, 2021.

[29] Richard Hochschild. Improving the precision of biological age determinations. Part 1: a new approach to calculating biological age. *Experimental Gerontology*, 24(4):289–300, 1989.

[30] David H Meyer and Björn Schumacher. Aging clocks based on accumulating stochastic variation. *Nature Aging*, pages 1–15, 2024.

[31] M.A. Ikram. The use and misuse of 'biological aging' in health research. *Nature Medicine*, 30(11): 3045–3045, 2024.

[32] Jarod Rutledge, Hamilton Oh, and Tony Wyss-Coray. Measuring biological age using omics data. *Nature Reviews Genetics*, pages 1–13, 2022.

[33] M Austin Argentieri, Sihao Xiao, Derrick Bennett, Laura Winchester, Alejo J Nevado-Holgado, Upamanyu Ghose, Ashwag Albukhari, Pang Yao, Mohsen Mazidi, Jun Lv, et al. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nature Medicine*, 30(9):2450–2460, 2024.

[34] Julian Mutz, Raquel Iniesta, and Cathryn M Lewis. Metabolomic age (MileAge) predicts health and life span: A comparison of multiple machine learning algorithms. *Science Advances*, 10(51):eadp3743, 2024.

[35] Ludger JE Goeminne, Anastasiya Vladimirova, Alec Eames, Alexander Tyshkovskiy, M Austin Argentieri, Kejun Ying, Mahdi Moqri, and Vadim N Gladyshev. Plasma protein-based organ-specific aging and mortality models unveil diseases as accelerated aging of organismal systems. *Cell Metabolism*, 37(1):205–222, 2025.

[36] Haidong Wang, Katherine R Paulson, Spencer A Pease, Stefanie Watson, Haley Comfort, Peng Zheng, Aleksandr Y Aravkin, Catherine Bisignano, Ryan M Barber, Tahiya Alam, et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *The Lancet*, 399(10334):1513–1536, 2022.

[37] Nazrul Islam, Vladimir M Shkolnikov, Rolando J Acosta, Ilya Klimkin, Ichiro Kawachi, Rafael A Irizarry, Gianfranco Alicandro, Kamlesh Khunti, Tom Yates, Dmitri A Jdanov, et al. Excess deaths associated with Covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *BMJ*, 373, 2021.

[38] Edgardo R Sepulveda, Nathan M Stall, and Samir K Sinha. A comparison of COVID-19 mortality rates among long-term care residents in 12 OECD countries. *Journal of the American Medical Directors Association*, 21(11):1572–1574, 2020.

[39] Marlies Bär, Judith AM Bom, Pieter LH Bakx, Cees MPM Hertogh, and Bram Wouterse. Variation in excess mortality across nursing homes in the Netherlands during the COVID-19 pandemic. *Journal of the American Medical Directors Association*, 25(9):105116, 2024.