
Article:
Davillas, Apostolos and Jones, Andrew Michael orcid.org/0000-0003-4114-1785 (2025) Biological age and predicting future health care utilisation. Journal of Health Economics. 102956. ISSN 0167-6296

https://doi.org/10.1016/j.jhealeco.2024.102956

Journal of Health Economics 99 (2025) 102956

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhealeco

Biological age and predicting future health care utilisation☆

Apostolos Davillas [a,b], Andrew M. Jones [c,*]

[a] Department of Economics, University of Macedonia, Bonn, IZA, Greece
[b] IZA Bonn, Germany
[c] Department of Economics and Related Studies, University of York, York YO10 5DD, UK

ARTICLE INFO

JEL classification:
C5
C81
I10
I18

Keywords:
Epigenetics
Biological age
Health care utilisation
Red herring hypothesis
LASSO
Supervised machine learning

ABSTRACT

We explore the role of epigenetic biological age in predicting subsequent health care utilisation. We use longitudinal data from the UK Understanding Society panel, capitalising on the availability of baseline epigenetic biological age measures along with data on general practitioner (GP) consultations, outpatient (OP) visits, and hospital inpatient (IP) care collected 5–12 years from baseline. Using least absolute shrinkage and selection operator (LASSO) regression analyses and accounting for participants' pre-existing health conditions, baseline biological underlying health, and socio-economic predictors we find that biological age is selected as a predictor of future GP consultations and IP care, while chronological rather than biological age is selected for future OP visits. Post-selection prediction analysis and Shapley-Shorrocks decompositions, comparing our preferred prediction models to models that replace biological age with chronological age, suggest that biological ageing has a stronger role in the models predicting future IP care as opposed to "gatekeeping" GP consultations.

1. Introduction

Much of the existing research on the individual-level determinants of health care demand, and their consequent costs, focuses on the role of age, morbidity and time-to-death (e.g., Brilleman et al., 2014; Carreras et al., 2018; Costa-Font and Vilaplana-Prieto, 2020; De Meijer et al., 2011; Howdon and Rice, 2018; Zweifel et al., 1999). Given that the proportion of those aged 65 and over is projected to continue increasing across OECD countries (OECD, 2021), a better understanding of the role of ageing as a predictor of health care utilisation is of particular importance. However, there is an unresolved debate over whether the effect of ageing on health care utilisation and expenditure may be overestimated. According to the "red herring" hypothesis, proximity to death rather than chronological age is a driver of healthcare expenditure (Zweifel et al., 1999). Potential omitted variables biases, that result from partially adjusted or unadjusted associations between morbidity and an individual's chronological age, may be a source of the over-estimation of the role of ageing on health care utilisation and expenditure (Costa-Font and Vilaplana-Prieto, 2020). For these reasons, relying on

☆ Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. The funders, data creators and UK Data Service have no responsibility for the contents of this paper. We are grateful for comments from Jason Fletcher, Erzo Luttmer, John Mullahy and participants at the Empirical Health Economics Workshop, Pisa, July 2024, and two anonymous reviewers. For the purpose of open access, a Creative Commons Attribution (CC BY) licence is applied to any Author Accepted Manuscript version arising from this submission.

* Corresponding author.
E-mail address: andrew.jones@york.ac.uk (A.M. Jones).

https://doi.org/10.1016/j.jhealeco.2024.102956
Received 19 July 2024; Received in revised form 29 October 2024; Accepted 5 December 2024
Available online 6 December 2024

simple projections based on chronological age does not necessarily provide an accurate picture of future health care utilisation and expenditure (Davillas and Pudney, 2020a, 2020b).

Many of the existing studies explore the associations between morbidity profiles and health care utilisation and costs either during short time horizons or focusing solely on patients who used health care during a limited time interval, or on those individuals registered with specific health insurers rather than the general population (Brilleman et al., 2014; Carreras et al., 2018; Howdon and Rice, 2018; Zweifel et al., 1999). Davillas and Pudney (2020a, 2020b) explore the role of biomarkers as predictors of future disability, healthcare demand and the subsequent costs. They use biomarkers separately, or as a composite measure (allostatic load), to proxy underlying health, with data from Understanding Society – the UK Household Longitudinal Study (UKHLS).

An important development in research based on large-scale social surveys, such as the UKHLS (Bao et al., 2022) and the Health and Retirement Study in the U.S.A. (Crimmins et al., 2020), is the integration of epigenetic biological age measures that are markers of the fundamental underlying mechanisms of healthy ageing (Horvath and Raj, 2018). Unlike chronological age, which universally progresses at the same rate, individuals can experience a higher or lower biological age than their chronological age. Epigenetic biological age is a strong predictor of all-cause mortality above and beyond chronological age, even after adjusting for traditional risk factors, such as unhealthy lifestyles and morbidity profiles (e.g., Chen et al., 2016).

Our study explores the role of epigenetic biological ageing in predicting subsequent health care utilisation over a period of 5–12 years after the baseline; UKHLS follows up the respondents from baseline for 5–12 years, collecting data on general practitioner (GP) consultations, outpatient/day-patient (OP) visits, and hospital inpatient (IP) care. We assess whether biological age plays a role over and above underlying physical health status at baseline (proxied by a composite biomarker index – allostatic load), pre-existing health conditions diagnosed before baseline, chronological age, and a set of baseline socio-economic and demographic characteristics (income, education, etc.). We use longitudinal data from the UKHLS, capitalizing on the recent release of epigenetic biological age measures (Bao et al., 2022; Benzeval et al., 2023). These are based on DNA extracted from blood samples collected at nurse visits as part of Waves 2 and 3 of UKHLS (our baseline). Demographic and socioeconomic characteristics of the respondents, along with a detailed set of nurse-collected and blood-based biomarkers, are also available at baseline. The dataset collects predetermined initial health conditions as self-reports prior to baseline.

Given that proximity to death, which has been highlighted as a key predictor of health care utilisation and expenditure (e.g., Carreras et al., 2018; Costa-Font and Vilaplana-Prieto, 2020; Howdon and Rice, 2018), is unobserved prior to death, it is of limited practical use for projecting individual health care utilisation for the needs of health policy. However, building on Grossman's seminal work on demand for health demand (Grossman, 1972), biological ageing can be considered as a more proximal measure of cumulative adverse health exposures, depreciation over time and investments. Unlike other proxies of people's health, including clinical biomarkers, that are not sufficiently representative of the underlying ageing mechanisms, epigenetic biological ageing is particularly relevant for researching healthy ageing (Horvath and Raj, 2018). Epigenetic biological ageing reflects the interaction between genes and the environment, through reversible mechanisms that regulate the function of the genome as a response to environmental exposures and, thus, modulate the ageing process (Bafei and Shen, 2023; Horvath and Raj, 2018). For example, Dalgaard and Strulik (2014) develop an economic model of ageing in which age of death is determined by optimal health investments and is relevant to biological ageing – chronological ageing and mortality are inevitable but individuals can slow down biological ageing by investing in their health. Our analysis provides econometric evidence on the respective predictive roles of biological age and chronological age on future health care utilisation.

The fact that our predictive analysis for future health care utilisation allows us to account for both predetermined health conditions as well as current underlying health at baseline (using bio-measures) responds to concerns about the confounding role of morbidity in the existing literature of the association between ageing and health care utilisation or expenditure (Costa-Font and Vilaplana-Prieto, 2020; de Meijer et al., 2011). Moreover, most of the relevant studies (e.g., Brilleman et al., 2014; Carreras et al., 2018; Howdon and Rice, 2018; Zweifel et al., 1999) do not identify pre-symptomatic or pre-diagnosed individuals who are at elevated risk of future health care utilisation. However, forward-looking policies to improve the efficiency of the healthcare system through control and better allocation of healthcare services require information from the general population.

Using UKHLS allows us to account for demographic and socioeconomic confounding effects that are often limited in the health care system records, as well as having coverage of individuals with latent health conditions that have not reached clinical endpoints (such as diagnosis). Unlike clinical endpoints, biological ageing measures may better capture "pre-clinical ageing" (Levine et al., 2018) and, thus, they may provide information on those at risk for elevated future health care utilisation. Our study extends and contributes to the literature on the predictors of health care utilisation for the general population, including "apparently healthy" individuals who have not reached the stage of diagnosis (e.g., Davillas and Pudney, 2020a, 2020b). We provide novel evidence on the predictive role of epigenetic biological age over and above the role of well-established predictors of health care demand. Enhancing our understanding of the risk factors for future health care utilisation may help us to better characterize the profile of those who are likely to experience higher demand for health care services.

Improving the accuracy of predictions is of high relevance in the context of health care utilisation and expenditure. Providing new evidence on the risk factors for future health care utilisation may enhance our ability to characterize the profile of those at risk of needing more care. As predictions of future health care resource utilisation are instrumental for better allocation and management of medical resources (Cui et al., 2018), the need to obtain better, more accurate and generalisable in the context of new (real-world) data is important. Recent arguments have highlighted that supervised machine learning methods, such as LASSO, may be particularly useful in health economics research for identification of predictors of health outcomes as well as for accurate and generalizable predictive analytics of health outcomes (Padula et al., 2022).

We use least absolute shrinkage and selection operator (LASSO) regression analysis to assess whether the availability of measures of

epigenetic biological age enhances predictions of future health care utilisation. LASSO is a supervised machine learning algorithm that performs variable selection and regularization to enhance the accuracy and interpretability of the resulting predictive model of future health care utilisation. Specifically, we use separate LASSO predictive regressions for GP consultations, OP visits and IP care, observed over a horizon of 5–12 years from baseline.

Penalized regression, especially the LASSO, assists in predicting an outcome of interest by selecting the subset of the variables that minimizes (out-of-sample) prediction error. Specifically, the LASSO estimator minimises the out-of-sample prediction error, balancing bias and variance to build an accurate predictive model. In essence, LASSO selects the predictors to be included in the model such that the fitted model is suitable for making out-of-sample predictions. As a regularization technique LASSO aims to prevent overfitting and enhance predictive accuracy. In the seminal work of Tibshirani (1996), LASSO is motivated over OLS based on two major advantages. Due to the nature of the L1 norm constraint, LASSO sets some of the coefficient estimates exactly to zero and, thus, removes selected predictors from the model. As such, LASSO serves as a model selection technique and facilitates model interpretation (Ahrens et al., 2020; Tibshirani, 1996). Secondly, LASSO outperforms OLS in prediction accuracy. The selection of a more parsimonious model using LASSO techniques, improves model generalisation by limiting the risk of overfitting and enhanced performance, allows for a deeper insight into the underlying processes that generated the data and an increased probability of generalisability of the findings to new data (Padula et al., 2022).

We find that pre-existing chronic health conditions and/or current underlying health, proxied by allostatic load, contribute to the prediction of GP consultations, OP visits and IP care 5–12 years from baseline. However, in these predictive models, there are differences in the roles of biological and chronological age; biological rather than chronological age is selected as a predictor to GP consultations and IP care, while chronological rather than biological age is the selected predictor for OP visits. Post-selection prediction analysis for GP and IP care and Shapley-Shorrocks decompositions comparing our preferred prediction models to models that replace biological age with chronological age, reveal more pronounced differences between biological and chronological age in the case of IP care as opposed to GP models. This suggests that biological age has a stronger role in the IP care prediction models rather than "NHS gatekeeping" GP consultations.

2. Data

The UKHLS, known as Understanding Society, is a nationally representative longitudinal study, running continuously from the initial wave in 2009–10, with each panel member interviewed annually. Its predecessor, the British Household Panel Survey (BHPS) was incorporated into the UKHLS from Wave 2. A feature of UKHLS is the inclusion of biosocial data, collected at a nurse visit, that provides biomeasures, blood-based biomarkers, along with genetic and epigenetic markers (Benzeval et al., 2023). Physical health measures and non-fasted blood samples were collected at nurse visits, conducted on average five months after the main Wave 2 interview for the UKHLS and similarly after Wave 3 for the BHPS sample. Respondents were eligible for nurse visits and collection of blood samples if they were aged 16 or over, lived in England, Wales or Scotland, were not pregnant, had no clotting or bleeding disorders and no history of fits.[1]

We follow the pooled UKHLS and BHPS sub-sample (Waves 2 and 3 for the UKHLS and BHPS sub-samples, respectively) with non-missing (nurse-collected and blood-based) biomarker data and contemporaneous socioeconomic characteristics up to Waves 7–13, where subsequent health care utilisation measures are collected. A selected set of biomarkers is used to construct allostatic load in order to proxy the underlying biological health of respondents at baseline (Waves 2/3). Our potential sample is restricted to those for whom the epigenetic measure of biological ageing is recorded (3,654 individuals in total) — a sub-sample of the full pooled wave 2 (UKHLS sample) and wave 3 (BHPS sample) sample that is restricted (by survey design) to: a) those for which nurse visits are conducted after the corresponding main waves 2/3; b) blood samples are taken; and c) epigenetic measure of biological ageing are available[2]. This potential sample with valid biological ageing data is restricted to 2,443 individuals, after excluding missing data on allostatic load and socioeconomic characteristics at baseline[3]. Our working sample is further restricted to follow-up respondents at UKHLS waves 7–13, where measures of health care utilisation are collected. This results in a working sample ranging between 1,339 and 1,632 observations depending on the measure of health care utilisation that is used.

To provide evidence on the potential implications of these sample restrictions, comparisons of the mean values of our predictors between the full pooled sample for Waves 2/3 (pooled UKHLS/Wave 2 and BHPS/Wave 3 sample that is not restricted to those selected for nurse visits and blood sample collection; 43,528 observations), and the maximum estimation sample (1,632 observations) are

[1] Participants gave informed written consent for their blood to be taken and stored for future scientific analysis. Nurse data collection at UKHLS has been approved by the National Research Ethics Service (10/H0604/2).

[2] From the sample of those with valid epigenetic measures of biological age (3,654 individuals in total), 131 individuals (about 3.6% of the epigenetic data) die between baseline and all subsequent waves (Waves 4-13). 67 of these deaths occurred during the Covid-19 outbreak. Linear probability models for attrition due to mortality show that biological ageing does not have a statistically significant association (p-value= 0.107) after adjusting for chronological age, gender and socioeconomic status. On the other hand, chronological age does have a statistically significant association (in mortality attrition models that both include and exclude the COVID-19 period). Linear probability models of attrition due to non-mortality illness-health reasons (during all subsequent waves from baseline) show that neither biological nor chronological age are statistically significant (after adjusting for gender and irrespective of whether or not individuals' socioeconomic status is accounted for). Our analysis is restricted to the survivors.

[3] The availability of epigenetic data in UKHLS is limited to those of white European origin.

available in Table A.1, in the Appendix. Despite the considerable reduction in the sample size between the two samples, the mean values of our predictors are similar between the two samples. This suggests that the restrictions imposed on our estimation sample (valid data on biomarkers and biological age, non-missing explanatory covariates as well as successive follow-up of respondents at subsequent waves) may have a limited impact on the composition of our estimation sample compared to the unrestricted full sample that represents a (two-stage stratified) random sample of the general population in Great Britain.[4]

Health care utilisation

Retrospective measures of health care utilisation, for the 12 months prior to interview, are collected at each UKHLS wave between 7 and 13. UKHLS collects data on the numbers of: GP consultations, attendances at a hospital or clinic as an out-patient or day patient (OP), and hospital in-patient (IP) days. The number of GP consultations and OP visits in the 12 months prior to interview (Waves 7–13) are collected into five intervals (0, 1–2, 3–5, 6–10, more than 10). Regarding the IP measures (Waves 7–13), respondents reported the number of days spent in a hospital or clinic as an inpatient in the same period.

We create measures of cumulative health care utilisation that cover the whole period between Waves 7 and 13. Specifically, for each of the GP and OP measures (Waves 7–13) we:

a) recode the reported five categories regarding the number of consultations/visits as 0, 1.5, 4, 8 and 10, reflecting the mid point for each reported interval, while keeping the first category as zero visits/consultations and top coding the last category to 10, and;
b) sum the resulting measures across waves 7–13 to create one cumulative measure for GP consultations and one for OP visits. For IP care, we use the IP measures described above to create a dichotomous measure that takes the value of one if the respondent experienced any IP days over the whole period between Waves 7–13, and zero otherwise.

Over and above these cumulative measures of health care utilisation, we also use measures that capture the longest time horizon from baseline; these measures are based on health care utilisation data from UKHLS Wave 13. Specifically, we create separate dichotomous variables on whether respondents had any GP consultation, OP visits or IP days at Wave 13. Summary statistics for all health care utilisation outcomes are shown in Table 1.

Biological and chronological age

DNA methylation-based measures, often called “epigenetic age” or DNA methylation age (“DNAmAge”), have been shown to be robust biomarkers of biological ageing (Jylhävä et al., 2017; Horvath and Raj, 2018). Methylation is a mechanism that drives human ageing and varies across people of the same chronological age (Fransquet et al., 2019). Using algorithm-derived weighted averages of methylation levels at different regions in the DNA sequence, so-called “epigenetic clocks” produce estimates of biological age.

Although chronological and biological age are correlated, chronological age measures the time since birth while biological age reflects the epigenetic interaction of genes and environment, with DNA methylation influencing the decline in viability of bodily organs over time (Bafei and Shen, 2023). Whereas chronological age increases at the same rate for everyone, biological age does not, with some people experiencing a higher or lower biological age than their chronological age. It has been shown that those of higher biological age, compared to chronological age or after adjusting for chronological age, may experience higher mortality and morbidity risks, functional limitations, and cognitive dysfunction (Chen et al., 2016; Faul et al., 2023; Li et al., 2022). Unlike many clinical biomarkers, that are not sufficiently representative of the fundamental underlying mechanisms of ageing, markers for biological ageing are particularly relevant for researching healthy ageing (Horvath and Raj, 2018).[5]

Recent UKHLS data releases provide epigenetic clocks (as individual-level derived variables) based on DNA methylation profiles that are extracted from blood samples collected at nurse visits as part of Waves 2 and 3. This methylation analysis is done for a sub-sample of those from whom blood data are collected and who consent to genetic analysis of their blood data (Bao et al., 2022). As we are interested in capturing biological age, in addition to chronological age, we use the epigenetic biomarker of ageing proposed by Levine et al. (2018), often called DNAm “PhenoAge” (Wang et al., 2023). As one of the leading second-generation epigenetic measures, DNAm PhenoAge has been developed to predict risks for all-cause mortality and can capture the process of ageing beyond chronological age, functioning as a marker of biological ageing (Zavala et al., 2024). PhenoAge captures risk across multiple tissues and outperforms alternative biological ageing proxies, strongly predicting a variety of ageing outcomes such as all-cause mortality,

[4] As an additional way to evaluate the potential implications of restricting our sample to valid epigenetic biological ageing data per se, Table A.2 in the Appendix provides comparisons of descriptive statistics between our maximum estimation sample and a comparison sample on which we have imposed the same restrictions as our estimation sample without conditioning on having valid biological age data. Very little difference is observed in the mean values between the two samples; this suggests that conditioning on valid epigenetic biological ageing per se (over and above the restrictions of valid data on biomarker data, non-missing explanatory covariates as well as successive follow-ups of respondents at subsequent waves) may have limited impact.

[5] For example, McCrory et al. (2020) argue that epigenetic biological age measures that are independent of chronological age (i.e., biological age measures from which chronological age is subtracted from them to create measures of epigenetic age acceleration) are either non-systematically or only weakly (correlation coefficient of 0.21) correlated with allostatic load — a composite measure of clinical biomarkers (which is described in the sub-section below and is used as a predictor to account for respondents’ baseline health in our prediction models).

Table 1
Summary statistics for the outcomes and predictors.

	Mean	Obs.
Health care utilisation outcomes		
GP visits (Waves 7–13)	15.590	1,372
Outpatient visits (Waves 7–13)	10.083	1,373
Any inpatient day (Waves 7–13)†	0.302	1,339
Any GP visit (Wave 13)†	0.580	1,630
Any outpatient visit (Wave 13)†	0.394	1,632
Any inpatient day (Wave 13)†	0.065	1,626
Predictors		
Initial diagnosed health condition: none†	0.642	1,632
Initial diagnosed health condition: present†	0.358	1,632
Allostatic load	-0.491	1,632
Chronological age	52.489	1,632
Biological age	44.322	1,632
Female†	0.563	1,632
Male†	0.437	1,632
Log household income	7.382	1,632
Secondary/above education†	0.830	1,632
No/basic education†	0.170	1,632
Non rented home†	0.862	1,632
Rent home†	0.138	1,632
England†	0.820	1,632
Wales†	0.086	1,632
Scotland†	0.094	1,632
Rural†	0.290	1,632
Urban†	0.710	1,632
Wave 2†	0.618	1,632
Wave 3†	0.382	1,632

† Dichotomous variable.

Note: Mean values for the explanatory variables are calculated for the maximum estimation sample size; the exact sample size for each model depends on item missingness for the healthcare utilisation variable.

cancers, and physical functioning (Levine et al., 2018).

Apart from PhenoAge UKHLS provides first-generation epigenetic measures, such as the "Horvath" (Horvath, 2013) and the "Hannum" (Hannum et al., 2013), the "Lin" (Lin et al., 2016) second generation measure as well as for the "DunedinPoAm" epigenetic measure, which reflects the 'rate' of biological ageing rather than the level *per se* (Wang et al., 2023; Benzeval et al., 2023). As a sensitivity analysis, regarding the selection of the biological ageing measure, we employ LASSO models that account for all biological ageing measures described above (Horvath, Hannum, Lin, DunedinPoAm and PhenoAge). This allows LASSO to select the biological ageing measure(s) to be used as predictors of future health care utilisation measures. The relevant results (described in the sensitivity analysis of our study below) confirm our choice of PhenoAge as the main measure of epigenetic biological age. This is broadly in line with existing literature that has shown that PhenoAge outperforms the Lin biological age measure and the first-generation measures in regard to predictions for people's mortality and a set of morbidity risks (e.g., Levine et al., 2018). These studies do not assess the performance of PhenoAge over alternatives in terms of predicting subsequent health care utilisation.

We also account for chronological age among the potential predictors of health care utilisation. Chronological age is a derived variable in the dataset; it is calculated in completed years from date of birth up to the nurse visits date, when the blood sample data are collected. Mean values for biological and chronological age are shown in Table 1: the mean chronological age is 52.5 years, while the mean of PhenoAge is considerably lower (44.3 years) suggesting that – on average – for the respondents in our sample their body is ageing more slowly than is considered normal for their chronological age.

Allostatic load

Allostatic load is a composite measure of a set of nurse-collected and blood-based biomarkers, which gives an assessment of a respondent's physiological condition. Allostatic load is elevated when a person's biological systems are affected by repeated physical and psycho-social stressors and it is associated with increased morbidity and all-cause mortality risks (Parker et al., 2022).

We use allostatic load at baseline (Waves 2 and 3 for the UKHLS and BHPS data, respectively) to proxy respondents' physiological health status. Given that allostatic load captures chronic physiological dysregulation, relevant to chronic morbidity, its predictive role relates more to health care demand due to these chronic conditions, rather than transient infections and/or accidental injury (Davillas and Pudney, 2020a).

Following Davillas and Pudney (2020a), our allostatic load index combines markers for adiposity, blood pressure, heart rate, lung function, inflammation, blood sugar levels, cholesterol levels, liver function and a steroid hormone collected at nurse visits following Waves 2/3 for the UKHLS and BHPS samples. Specifically, the waist-to-height ratio is used to measure adiposity, while the resting heart rate, systolic blood pressure and high-density lipoprotein cholesterol (HDL) are used to measure cardiovascular health. Forced vital

capacity (FVC), the total amount of air forcibly blown out after a full inspiration, is used to measure lung function; higher FVC values indicate better lung functioning. C-reactive protein, an inflammatory biomarker, and glycated haemoglobin (HbA1c), a sugar in the blood biomarker which is a validated diagnostic test for diabetes, are included in our composite allostatic load measure. Liver functioning is captured in our allostatic load measures using albumin; low albumin levels suggesting impaired liver function. A steroid hormone (dihydroepiandrosterone sulphate; DHEAS), associated with cardiovascular risk and all-cause mortality through psycho-social mechanisms (Ohlsson et al., 2010), is also used. To calculate our composite allostatic load measure we convert HDL, FVC, Albumin and DHEAS to negative values to reflect ill-health rather than good health, and then transform each of the biomarkers into a z-score; these nine z-scores are summed to produce the composite measure.

Diagnosed health conditions

In our prediction models for health care utilisation, we also account for pre-existing diagnosed health conditions obtained from self-reports made before the baseline biomarker measurements (and biological ageing measures) as part of Waves 2 (for the UKHLS sample) and 3 (for the BHPS sample). Specifically, we use a dichotomous variable that takes the value of one if the individual reported any past diagnosis of a long-lasting health condition (asthma, chronic bronchitis, congestive heart failure, coronary heart disease, heart attack or myocardial infarction, stroke, cancer or malignancy, diabetes, high blood pressure, arthritis, and liver condition) before the baseline biomarker measurements were taken, and zero otherwise.

Other predictors

We include a set of additional predictors in our health care utilisation predictions which have been argued to be associated (directly or indirectly) with health outcomes (Carrieri and Jones, 2017; Davillas et al., 2019; Davillas and Pudney, 2020a, 2020b; Van Doorslaer et al., 2004). These variables are collected at baseline as part of the main Wave 2 (for the UKHLS sample) and Wave 3 (for the BHPS sample). In line with Davillas and Pudney (2020a), we include the following demographic and socio-economic variables in our prediction analysis.

We account for sex as a basic demographic variable in our predictive models. Moreover, a set of socioeconomic indicators are also used. Household income is the sum of the incomes of all household members. Household income is deflated, using monthly Retail Price Indexes, to facilitate comparisons over time (given the pooling of Wave 2 and 3 data for baseline measurements and within-wave variation in interview dates) and equivalized, using the modified OECD equivalence scale, to account for household composition; then, it is log transformed to allow for concavity of the health-income association. Educational attainment is captured using a dichotomous variable of whether respondents have completed secondary or post-secondary education as opposed to basic or no qualification. We also account for housing tenure (not living in a rented home versus a rented home). A dichotomous variable for living in an urban area and a set of dichotomous variables to capture the three nations of Great Britain (England, Scotland and Wales) are also included as health care systems policy is determined on a national basis (Davillas and Pudney, 2020a). Finally, we account for wave dummies to capture time effects as pooled data from Wave 2 and 3 are used. Mean values for all control variables are presented in Table 1 for the maximum estimation sample.

3. Methods

Our objective is to assess whether the availability of measures of biological age enhances predictions of future health care utilisation relative to chronological age and biomarkers (allostatic load), along with other baseline predictors. To do this we adopt a model selection approach based on penalized regressions. As the emphasis is on selecting a sparse set of predictors, we adopt the standard LASSO (least absolute shrinkage and selection operator) estimator. LASSO performs variable selection and regularization, enhancing the prediction accuracy of the selected model (Tibshirani, 1996; Hastie et al., 2015). Our interest is whether the selected models, for making predictions of health care utilisation, include biological age instead of or as well as chronological age, allostatic load, initial diagnosed chronic conditions and the other predictors.

Assume a linear model to predict future health care utilisation (y_i), for each individual i $(i = 1,\ 2...,\ N)$, using the set of potential predictors $(x_1,\ x_2,...x_p; j = 1,\ 2,\ ...p)$. Given the assumption of sparsity, LASSO minimizes the mean squared prediction error subject to the L1 norm constraint on the absolute parameter values, which penalizes the complexity of the model. Specifically, the LASSO estimator $\widetilde{\beta_\lambda}$ of $\boldsymbol{\beta}$ minimizes:

$$Q_\lambda(\beta) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - X_i'\beta\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{1}$$

where, $\lambda \geq 0$ is a penalty or tuning parameter. The vector $\boldsymbol{X}$ includes the potential predictors, that are standardized so that the selection of predictors does not depend on their measurement scales. Different values of λ lead to different LASSO estimates; the penalty has the effect of forcing some of the coefficient estimates to be exactly equal to zero when the λ parameter is sufficiently large. LASSO minimizes the objective function (eq. 1) *for a grid* of values of λ. The algorithm chooses the solution that minimizes the out-of-sample prediction error based on 10-fold cross validation.[6]

Separate models are estimated to predict each of our measures of health care utilisation. In each model we include a set of potential predictors for the algorithm to select from: allostatic load, initial diagnosed health condition, biological age, chronological age, along with our set of socioeconomic and demographic controls (as described in the Data sub-section); we also include interaction terms between allostatic load and gender to capture gender differences in the predictive role of allostatic load on health risks (e.g., Demirer et al., 2021) as well as polynomials of biological and chronological age to capture non-linearities in the association between these variables and subsequent health care utilisation. All our continuous predictors are transformed into z-scores to have a mean of zero and a standard deviation of one.

Finally, we compute post-selection predictions for the number of subsequent GP consultations and for the probability of IP care (i.e., for those health care utilisation outcomes for which LASSO models select biological age as one of the predictors). These post-selection predictions, estimated at different values of biological age (with biological age being standardized using z-scores), are based on unpenalized OLS estimates for the sub-set of predictors selected by the LASSO algorithm; these are unbiased and may have better out of sample performance (Belloni and Chernozhukov, 2013; Cameron and Trivedi, 2022). For comparison purposes we also estimate the corresponding predictions (at different chronological age values) obtained from GP and IP care models on which biological age is replaced with chronological age (expressed in z-scores) in the list of predictors (with all remaining predictors remained the same). The two sets of predictions are compared graphically. Finally, Shapley-Shorrocks decompositions of the R-squared are computed and compared for both specifications (with z-scores of the chronological or biological age) for GP and IP care (Shorrocks, 2013).

4. Results

4.1. Main results

In this section, we present results from our base-case prediction model for each of the future health care utilisation outcomes, using PhenoAge as our epigenetic biological age measure. In sub-section 4.2 we conduct sensitivity analysis for the selection of the biological age measure in our analysis, confirming the choice of using PhenoAge as our main biological age measure. Fig. 1 presents cross-validation plots for the models predicting future GP, OP and IP care. These graphs provide an illustration that, in our prediction models for future health care utilisation, cross-validation chooses the model that minimizes the CV mean prediction error over the searching grid for the penalty parameter λ. For our cumulative GP consultations prediction model, the selected λ (λ_{CV}) that gives the minimum is λ_{CV}=0.18 (corresponding to a model with nine selected predictors); specifically, the CV mean prediction error decreases as the penalty λ decreases until λ_{CV}=0.18, after which it increases again reflecting the trade-off between bias and precision. In case of the dichotomous outcome for any GP consultation at wave 13, Fig. 1 (first row, right-hand side graph) shows that λ_{CV}=0.0041 (corresponding to a model with 9 predictors) is the λ that minimises the CV mean prediction error.

Turning to OP visits (Fig. 1, second row), the λ that minimizes the CV mean prediction error is λ_{CV}=0.082 (corresponding to a model with 13 selected predictors) for the case of the cumulative OP visits outcome. For the prediction model of any future OP visit for those followed up to the longest time interval from baseline ("Any outpatient visit (Wave 13)"), the selected lambda is λ_{CV}=0.01, corresponding to a model with six predictors. For the cumulative IP days ("Any inpatient day (Waves 7–13")) the selected parameter lambda is λ_{CV}=0.013, corresponding to a model with six selected predictors (Fig. 1, third row, left-hand side graph). While λ_{CV}=0.0081 (corresponding to a model with three predictors) for the probability of any IP day when individuals are followed up to the longest time horizon (Fig. 1, third row, right-hand side graph).

Tables 2-7 provide a summary of the values of lambda at which predictors are selected (knots) and the corresponding CV mean prediction errors for the case of our prediction models of future health care utilisation. The lambda parameter that minimizes the CV mean prediction error (λ_{CV}) is presented in the tables (signposted by a star) irrespective of whether it corresponds to a knot (i.e., to a lambda value at which predictors are selected); two knots that come after the λ_{CV} are also included in these tables for comparison purposes, confirming that the CV mean prediction error increases for lambdas lower than the λ_{CV}.

For GP consultations (Tables 2 and 3), allostatic load, proxying the health status of the respondents at baseline, as well as the (absence) of initial health conditions are the predictors that are selected in the first knots. Along with the main allostatic load predictor, the "Male*Allostatic load" interaction is selected as a predictor by LASSO suggesting a differential predictive role of allostatic load by sex. Of particular interest, biological rather than chronological age at baseline is selected as a predictor for future GP consultations; overall, this highlights the predictive power of biological, rather than chronological age, in predicting future GP services utilisation after accounting for health conditions and allostatic load, along with a set of demographic and socio-economic predictors (which

[6] Specifically, k-fold cross-validation randomly divides the data into k partitions/sub-samples, called folds. For each fold of the data a regression is fitted on the other nine folds, using the variables selected by LASSO for the given value of λ, and the mean squared error (MSE) is computed for that fold. These MSEs are averaged to give the CV mean prediction error. CV stops when the minimum of the CV function is found, and it sets the selected λ_{CV} to the λ that gives the minimum.

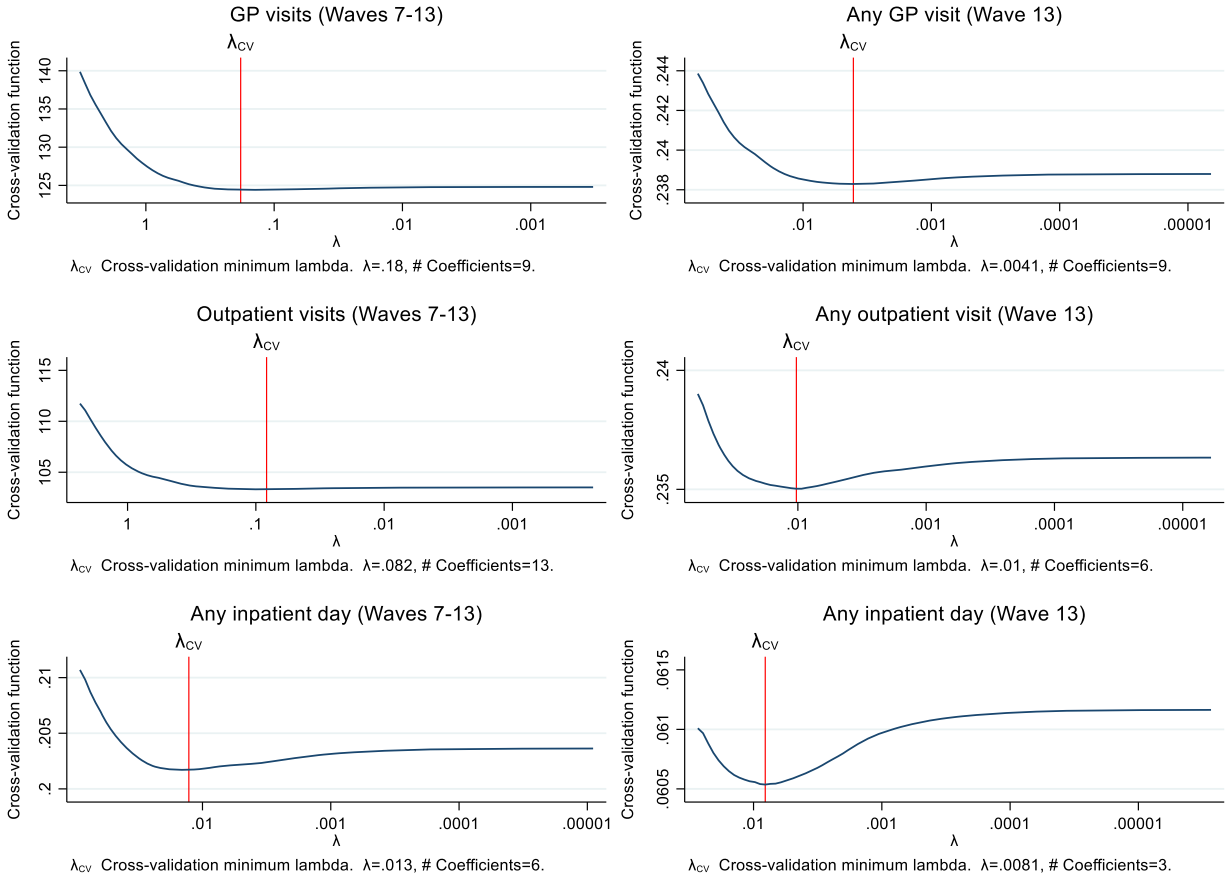

Fig. 1. Cross-validation function over the search grid for the penalty parameter lambda: GP, OP and IP care.

Table 2
CV mean prediction error and selected predictors across knots: GP visits (Waves 7–13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
2.973	1	138.3288	Allostatic load
2.249	2	134.4224	Initial diagnosed health condition: none
1.287	3	129.1370	Male
1.068	4	127.9497	Not rented home
0.808	5	126.5899	Secondary/above education
0.557	8	125.6055	Biological age; Male*Allostatic load; Wave 2
0.241	9	124.4653	Scotland
***0.182**	**9**	**124.4273**	
0.151	10	124.4330	Log household income
0.126	11	124.4465	Rural

Note: The parameters in bold (*) correspond to the lambda selected by cross-validation. Estimation sample size: 1,372.

include education, house tenure, income, urbanization and regional dummies, selected by LASSO depending on the GP outcome).

For OP care (Tables 4 and 5), the underlying health state at baseline, proxied by allostatic load, and the (absence) of initial health conditions contribute to the prediction of future OP health care utilisation. As for GP consultations, allostatic load exerts a different predictive role by sex. Over and above these predictors, chronological rather than biological age at baseline matters for subsequent OP care 5–12 years later. Unlike GP consultations, chronological age at baseline (and its squared term for the case of the cumulative OP outcome) is selected among the first predictors by LASSO in the case of OP visits models, while biological age is not selected at knots up to the minimum of the CV mean prediction error. Moreover, a wider set of baseline demographic and socioeconomic predictors are selected by LASSO in the case of the cumulative OP measure (Table 4) as opposed to the predicted probability of any OP visit 12 years later (Table 5).

Tables 6 and 7 present the results for IP care. Overall, these results show that baseline biological age and its squared term, current

Table 3
CV mean prediction error and selected predictors across knots: any GP visit (Wave 13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
0.0602	1	0.24339	Allostatic load
0.0500	2	0.24236	Initial diagnosed health condition: none
0.0345	3	0.24065	Male*Allostatic load
0.0314	4	0.24037	Biological age
0.0261	5	0.23997	Rural
0.0238	6	0.23980	Male
0.0136	7	0.23877	Not rented home
0.0113	8	0.23859	Log household income
0.0094	9	0.23849	England
***0.0041**	**9**	**0.23829**	
0.0025	10	0.23834	Chronological age
0.0009	11	0.23856	Wales

Note: The parameters in bold (*) correspond to the lambda selected by cross-validation. Estimation sample size: 1,630.

Table 4
CV mean prediction error and selected predictors across knots: OP visits (Waves 7–13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
2.1399	1	111.079	Allostatic load
1.9498	2	110.220	Chronological age
1.6187	3	108.587	Initial diagnosed health condition: none
0.5817	5	104.473	Chronological age squared; Not rented home
0.4830	7	104.206	Secondary/above education; Wave 2
0.3329	8	103.730	Wales
0.3033	9	103.651	Male*Allostatic load
0.2518	10	103.551	Male
0.2295	11	103.512	Log household income
0.1313	12	103.362	England
0.1090	13	103.337	Rural
***0.0825**	**13**	**103.325**	
0.0568	14	103.361	Biological age
0.0224	14	103.443	

Note: The parameters in bold (*) correspond to the lambda selected by cross-validation. Estimation sample size: 1,373.

Table 5
CV mean prediction error and selected predictors across knots: any OP visit (Wave 13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
0.0548	2	0.23853	Chronological age; Male*Allostatic load
0.0344	3	0.23617	Allostatic load
0.0237	4	0.23545	Initial diagnosed health condition: none
0.0180	5	0.23524	Male
0.0149	6	0.23515	Rural
***0.0103**	**6**	**0.23503**	
0.0078	7	0.23510	Biological age
0.0065	8	0.23519	Not rented home

Note: The parameters in bold correspond to the lambda selected by cross-validation. Estimation sample size: 1,632.

health (allostatic load) and (absence) of existing health conditions are predictors of IP care 5–12 years later on. On the other hand, chronological age at baseline is not selected as a predictor of IP care. Moreover, unlike GP and OP visits, the LASSO estimator results in much more parsimonious specifications for IP care as only education and household income (if any) are selected from our set of demographic and socio-economic predictors.

Fig. 2 presents the penalized LASSO coefficient estimates for the selected variables of interest (i.e., allostatic load, chronological age, biological age and initial health conditions — depending on whether they are relevant to each particular model). Overall, the coefficient signs are as expected. The absence of initial diagnosed health conditions has a negative sign for all GP, OP and IP prediction models. Allostatic load, with higher values indicating worse underlying health at baseline, has a positive sign across all our health care prediction models; interaction terms with gender show that the predictive role of allostatic load at baseline is more pronounced for males in the case of GP and OP visits 5–12 years later. Higher biological age at baseline positively predicts both our GP consultation outcomes 5–12 years later; non-linearities are observed for our IP outcomes, with the square polynomials in biological age showing a

Table 6
CV mean prediction error and selected predictors across knots: any inpatient day (Waves 7–13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
0.08211	1	0.209828	Biological age
0.06817	2	0.207758	Allostatic load
0.04698	3	0.204714	Secondary/above education
0.03554	4	0.203257	Initial diagnosed health condition: none
0.01853	5	0.201778	Biological age squared
0.01689	6	0.201741	Log household income
***0.01277**	**6**	**0.201721**	
0.01164	7	0.201734	Rural
0.00802	8	0.201961	Wave 2

Note: The parameters in bold (*) correspond to the lambda selected by cross-validation. Estimation sample size: 1,339.

Table 7
CV mean prediction error and selected predictors across knots: any inpatient day (Wave 13).

Lambda	No. of nonzero coef.	CV mean prediction error	Selected predictors
0.0247	1	0.060968	Biological age
0.0171	2	0.060689	Allostatic load
0.0107	3	0.060564	Biological age squared
***0.0081**	**3**	**0.060537**	
0.0061	4	0.060555	Male
0.0056	5	0.060567	Wales

Note: The parameters in bold (*) correspond to the lambda selected by cross-validation. Estimation sample size: 1,626.

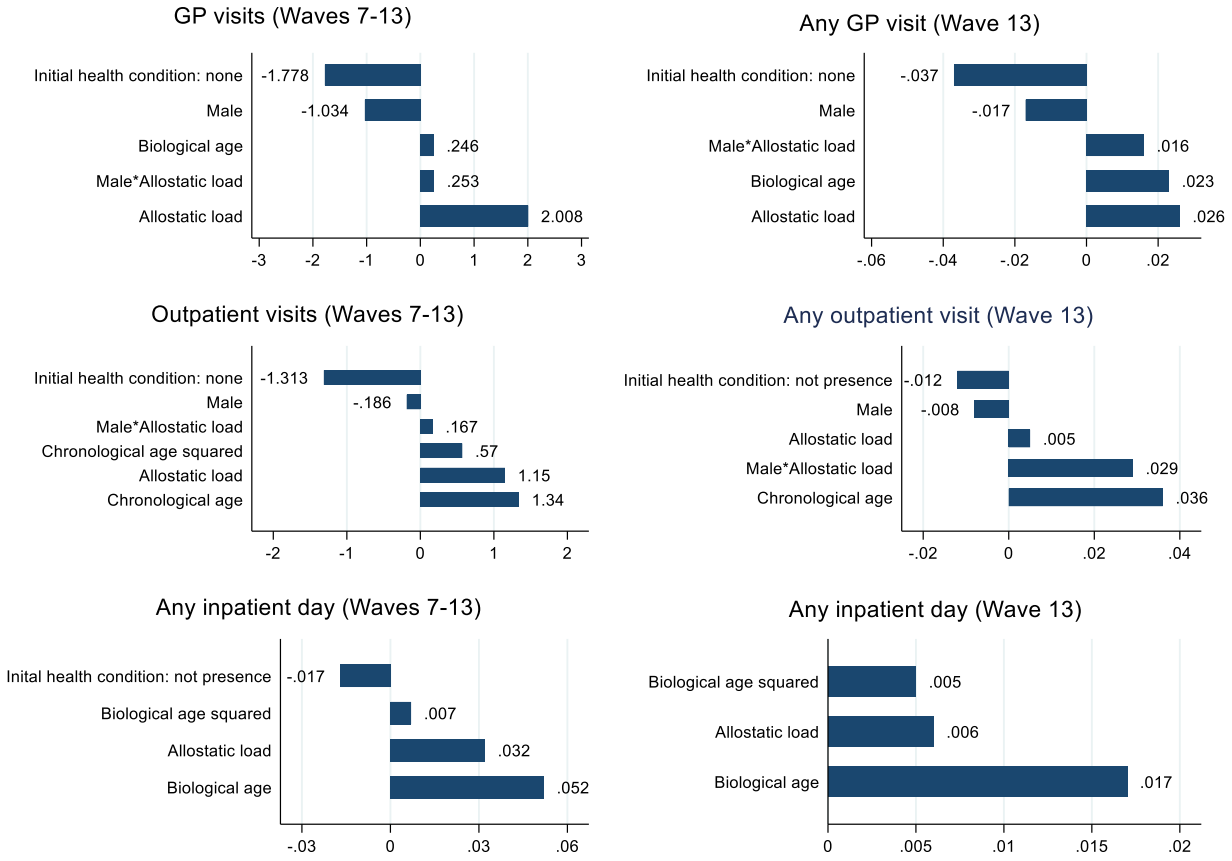

Fig. 2. Penalized coefficients for the selected set of (standardized) predictors: GP, OP and IP care.

stronger predictive role for those who are biologically older at baseline. Finally, the square polynomials in chronological age for OP care models suggest that chronological age at baseline exerts a stronger predictive role for subsequent OP care utilisation 5–12 years later, the older someone is at baseline.

The results so far show that biological, rather than chronological, age at baseline is selected for GP consultations and IP care 5–12 years later. On the other hand, chronological rather than biological age at baseline is selected for the prediction of future OP care. To show the predictive role of biological age for GP and IP care, post-selection predictions are computed using both biological age and also when the prediction models are re-estimated replacing biological with chronological age. The pattern of these predictions obtained at different levels of biological or chronological age (depending on whether z-scores of chronological or biological age is included as predictor) are compared in Fig. 3. It seems that the divergence between the predictions at z-scores of biological (blue lines) and chronological age distribution (red lines) are more pronounced for IP care at the right tails.

Table 8 compares the Shapley decompositions of the R-squared from the prediction models (with biological or chronological age) used to obtain the predictions in Fig. 3. For GP consultations, allostatic load and initial health conditions make the dominant contributions; estimating the corresponding prediction models that replace biological age with chronological age has little impact on the overall decomposition. However, for IP care, biological age is the dominant factor, accounting for about 43 per cent of the explained variation; when IP prediction models are estimated with biological age replaced with chronological age, this share declines (to 33 per cent) and there is a corresponding increase in the share of allostatic load. This suggests that, in terms of predicting IP care, biological age does a better job of picking up the underlying morbidity than chronological age.

4.2. Robustness checks

As a sensitivity analysis, regarding the selection of the biological ageing measure, we include all of the biological ageing measures described above (Horvath, Hannum, Lin, DunedinPoAm and PhenoAge) in the pool of potential predictors. Specifically, we estimate separate LASSO models for each of our cumulative GP consultations, OP visits and IP days including all five biological ageing measures (Horvath, Hannum, Lin, DunedinPoAm, and PhenoAge) along with the covariates including in our study. Table A.3 (Appendix) provides a summary table for the selected predictors (listed in order of selection at different knots) from LASSO models accounting for our full set of five epigenetic biological age measures. Overall, sensitivity analysis results confirm the use of PhenoAge as the main biological age measures in our analysis; specifically, PhenoAge is the selected biological age measure (from all five different biological age measures accounted for) for LASSO models that minimize the out-of-sample prediction errors for the cumulative GP consultations and IP days (i.e., the health care utilisation measures for which base-case results show that that biological age is a selected predictor among others).

These results are consistent with the existing literature as PhenoAge is a second-generation biological age measure that captures ageing process beyond chronological age, the different mortality and morbidly risks among individuals of the same chronological age and, as such, considered a biomarker of biological ageing (Zavala et al., 2024). Moreover, PhenoAge captures risk across multiple tissues and outperforms the Lin biological age measure and the first generation (Horvath and Hannum), which unlike the second-generation measures are trained to solely predict chronological age (Levine et al., 2018)[7].

We conduct sensitivity analysis for the robustness of our findings to the number of folds used for the K-fold cross-validation. Typically, a larger number of folds implies that the training set size increases so bias decreases; at the same time however, the fitted models are more likely to overlap and, thus, the test set predictions are more highly correlated leading to greater variance in the estimate of the expected prediction error. It has been argued that the most commonly used in empirical research 10-fold (as in our base-case analysis) provides a good balance between bias and variance (Cameron and Trivedi, 2022). Robustness checks estimating LASSO models using 5 or 20 CV folds suggest no changes to our base-case results on whether or not the variables of interest (i.e., chronological age, biological age, initial health conditions and allostatic load) are selected for the prediction model; for comparison purposes with the base-case results in Tables 2-7, Tables A.4–A.6 (Appendix) provide summaries of the selected predictors (listed in order of selection at different knots) for the LASSO models that use 5 or 20 CV folds and minimize the out-of-sample prediction errors.

As an additional sensitivity analysis, we re-run our prediction models using adaptive LASSO, rather than CV, to select the tuning parameter. These models result in a more parsimonious set of predictors (in some cases) - this is expected because adaptive LASSO models typically have fewer nonzero coefficients (Zou, 2006). However, adaptive LASSO results in no changes in the selected set of predictors of interest (i.e., chronological age, biological age, initial health conditions and allostatic load) compared to the corresponding base-case CV LASSO models. Tables A.4 -A.6 (Appendix) provide summary tables for the selected predictors from the adaptive LASSO models. It should be noted that our main results remain based on CV LASSO models because they are widely employed when the goal is prediction - the main scope of the analysis in our study.

For consistency across all health care utilisation outcomes, our base-case models are estimated using linear specifications. Robustness checks using probit LASSO (for our dichotomous health care utilisation outcomes) and Poisson LASSO (for our cumulative number of GP or OP consultations) results in identical sets of selected predictors as the base-case linear LASSO models presented in the

[7] According to the existing literature PhenoAge outperforms the Lin biological age measure and the first-generation measures specifically in regard to predictions for people's mortality and a set of morbidity risks. These biomedical studies do not assess the performance of PhenoAge over alternatives in terms of predicting subsequent individual's health care utilisation. We believe that – given the data available to us – it is reassuring to *directly* test our base-case choice of PhenoAge over alternative biological ageing measures on models explicitly aiming to predict future health care utilization 5-12 years from baseline.

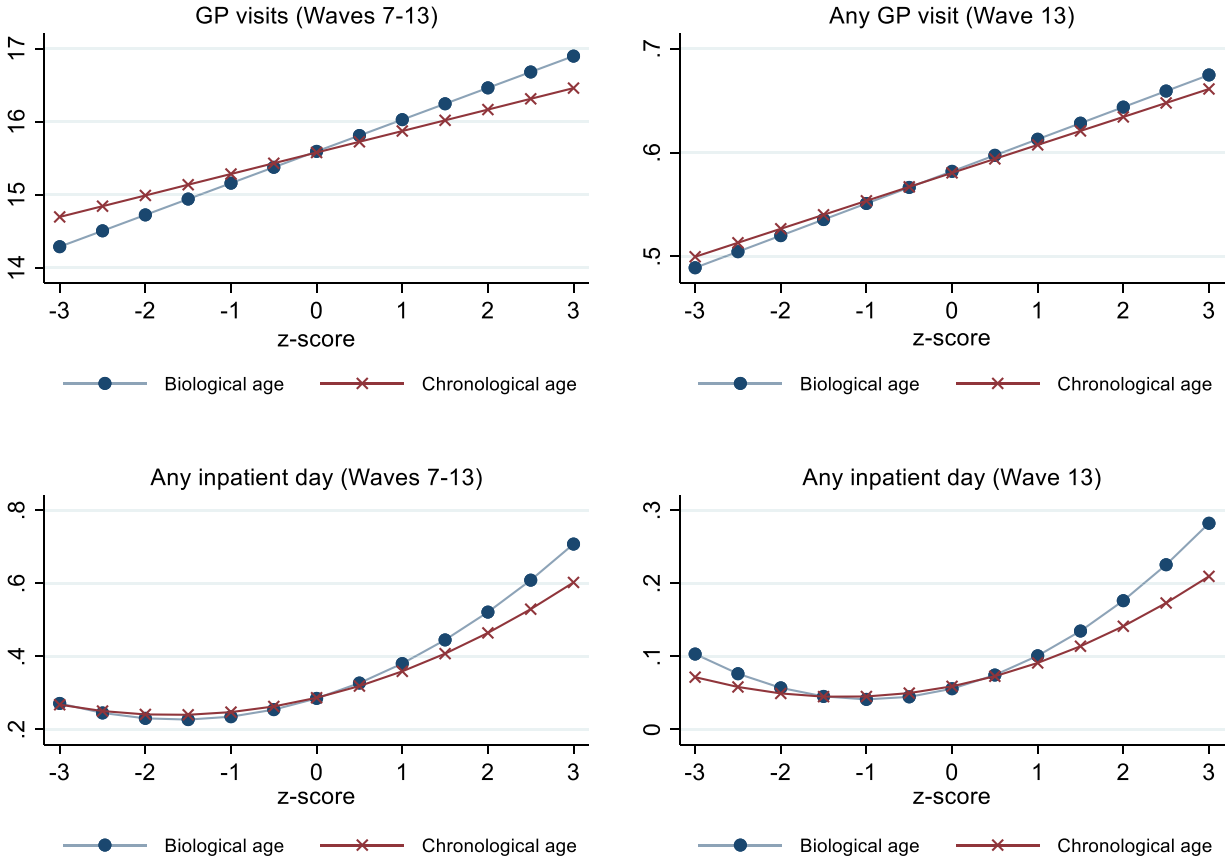

Fig. 3. Post-selection predictions across biological and chronological age: GP consultations and IP care.

Table 8
Shapley decomposition of R^2 for post-selection linear regressions (with biological age and when replaced with chronological age): GP visits and IP care.

	With biological age Percentage contribution	**With chronological age** Percentage contribution
Panel A. Number of GP visits (Waves 7–13)		
Age	8.44	6.87
Allostatic load	38.11	39.25
No initial health condition	26.48	26.81
Male	11.38	11.25
Secondary/above education	5.35	5.45
Not rented home	7.41	7.55
Scotland	0.57	0.55
Wave 2	2.26	2.27
Panel B. Probability of any inpatient day (Waves 7–13)		
Age	42.52	33.23
Allostatic load	25.58	30.36
No initial health condition	11.44	13.05
Secondary/above education	15.98	18.24
Log household income	4.48	5.12

study (Tables A.4-A.6, Poisson LASSO and Probit LASSO columns, as compared to Tables 2-7).

Our health care utilisation outcomes from Waves 7–13 include part of the COVID-19 period (2020–2022); to alleviate concerns on whether our prediction results are contaminated by the effect of the COVID-19 outbreak, we conduct sensitivity analysis using cumulative GP consultations, OP visits as well as measures on whether people experience any IP day based on Wave 7 (2015–2016)-Wave 10 (2018–2019) – a period that excludes the COVID-19 outbreak in the UK. The results of this sensitivity analysis (available upon request) are similar to our base-case results, further confirming the conclusions of our study.

Allostatic load is likely to be associated with both chronological and biological age. As a further sensitivity check we re-run our

analyses with allostatic load omitted from the set of potential predictors for the LASSO analysis (see Table A.7). The main conclusions regarding the respective roles of chronological and biological ageing across different health care utilization measures remained unchanged: biological rather than chronological age is selected as a predictor to GP consultations and IP care, while chronological rather than biological age is the selected predictor for OP visits.

5. Conclusion

In this study we use supervised machine learning techniques to explore whether biological age enhances our ability to predict future health care utilisation 5 to 12 years from baseline. Capitalising on recent advances in survey data incorporating epigenetic ageing measures, we study whether biological ageing plays a predictive role, over and above chronological age, current (allostatic load) and pre-existing underlying health status as well as baseline demographic and socio-economic correlates.

We find that pre-existing diagnosed chronic health conditions and/or current baseline underlying health, proxied by allostatic load, contribute to the prediction of health care demand for GP consultations, OP visits, and IP care. After accounting for these predictors along with baseline demographics and socio-economic factors, we find that biological rather than chronological age is selected as a predictor of GP consultations and IP care 5–12 years from baseline, while chronological and not biological age is selected as a predictor for future OP care. Post-selection prediction analysis and Shapley-Shorrocks decompositions for GP and IP care suggest that biological age has a stronger role in the IP care prediction models rather than "NHS gatekeeping" GP visits.

These results confirm exiting evidence on the role of baseline underlying health, proxied by allostatic load, to future health care utilisation (Davillas and Pudney, 2020a) and extend evidence on their predictive role when biological ageing measures are also accounted for. Our findings show that biological age can predict the more intense and severe cases that result in IP care as well as "NHS gatekeeping" GP consultations. On the other hand, the fact that chronological age is selected as a predictor of the typically more routine procedures (including diagnostics), which are delivered as OP care, may reflect dependency of the diagnostic referrals' decisions by patients chronological age (e.g., di Martino et al., 2023; Marx et al., 2022).

Our findings on the predictive role of biological ageing over and above underlying baseline health may be useful for tailored preventive interventions. Effective targeting of those with accelerated biological ageing could identify better the population groups with the highest potential of future healthcare needs and costs. This builds on existing evidence suggesting that epigenetic measures of biological ageing are considered valid proxies of healthy longevity (Phyo et al., 2024) and may be used for "risk stratification and risk estimation of future functional and cognitive aging" (Jain et al., 2022). In 2016, the US Food and Drug Administration approved the first epigenetic cancer test, which is efficient for the diagnosis of colorectal cancer for adults aged 50 years and older who may be asymptomatic and without a-priori excess risk of colorectal cancer (Dupras et al., 2019). Given the findings of our study, a wider use of epigenetic biological ageing measures on early diagnosis (or even prevention) of certain health conditions, especially for those that imply significant costs for society[8], may contribute towards a more efficient use of health care resources and on slowing the growth of healthcare costs.

Recent technical developments have expanded the options for obtaining epigenetic biological ageing measures with dried blood spots collected on filter paper, providing an accurate, low-cost and easy implemented alterative to the conventional venipuncture procedure (Ryan, 2021). To the extent that these technical improvements may mitigate a part of the administrative costs of epigenetic biological ageing measurement, our results may contribute to considerations of the cost-effectiveness of any potential pilot (or either feasibility) screening programmes based on biological ageing that target selected population groups.

Moreover, our findings on the independent predictive role of biological as opposed to chronological age may be relevant to the design of capitation-based payments that are used to allocate GP budgets in a number of countries (Brilleman et al., 2014, Sibley and Glazier, 2012, Shepherd, 2017). Developing capitation payments that more closely approximate the expected future health care utilisation in primary care, by including patient-level data on their underlying health as well as biological ageing, may reduce incentives for providers to engage in "cream-skimming". At least in the short-run, there are significant administrative costs along with growing ethical considerations of obtaining biological ageing measurements for all patients registered in a GP practice. However, the growing availability of large-scale multipurpose (nationally representative) survey data that incorporate epigenetic biological ageing measures may offer secondary-level anonymised data on the distribution of biological age at the regional level as an alternative; given the challenges associated with individual-level data on biological ageing, capitation payment formulas that incorporate regional-level variations in the distribution of biological age that are extracted from nationally representative survey data may better approximate actual future health care demand.

CRediT authorship contribution statement

Apostolos Davillas: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Andrew M. Jones:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization.

[8] For example, the total economic burden (including non-health care costs such as loss of productivity due to disability) of colorectal cancer across Europe in 2015 was €19,1 billion (Henderson et al., 2021).

Appendix: additional tables and figures

Table A.1
Summary statistics for the predictors: working sample versus the full original sample.

	Maximum working sample[†]		UKHLS (wave 2) & BHPS (wave 3) [‡‡]	
	Mean	Obs.	Mean	Obs.
Initial diagnosed health condition: none[†]	0.642	1,632	0.615	43,528
Initial diagnosed health condition: present[†]	0.358	1,632	0.385	43,528
Allostatic load[#]	-0.491	1,632	-0.355	8,334
Chronological age[##]	52.489	1,632	51.008	20,283
Biological age[###]	44.322	1,632	45.313	3,591
Female[†]	0.563	1,632	0.541	43,528
Male[†]	0.437	1,632	0.459	43,528
Log household income	7.382	1,632	7.242	43,528
Secondary/above education[†]	0.830	1,632	0.747	43,528
No/basic education[†]	0.170	1,632	0.253	43,528
Non rented home[†]	0.862	1,632	0.720	43,528
Rent home[†]	0.138	1,632	0.280	43,528
England[†]	0.820	1,632	0.803	43,528
Wales[†]	0.086	1,632	0.089	43,528
Scotland[†]	0.094	1,632	0.108	43,528
Rural[†]	0.290	1,632	0.250	43,528
Urban[†]	0.710	1,632	0.750	43,528
Wave 2[†]	0.618	1,632	0.787	43,528
Wave 3[†]	0.382	1,632	0.213	43,528

[†] Dichotomous variable.

[‡] Estimation sample as in Table 1; represents the pooled sample of Wave 2 (UKHLS)/ Wave 3 (BHPS) respondents that are followed up at subsequent waves, to obtain future health care use measures, and constrained having non-missing information on the explanatory covariates, on allostatic load, and on biological age. The exact estimation sample (for each health care utilisation regression model) depends on item missingness for the health care utilisation measures.

[‡‡] The full pooled sample of Wave 2 (UKHLS)/ Wave 3 (BHPS) respondents from which the restricted working samples are originated; these are pooled data from the main wave 2 (UKHLS sample) and wave 3 (BHPS sample) without imposing any of the restrictions as described above (including restriction to participate in the nurse visits that follow up the corresponding main waves).

[#] The mean of allostatic load for the full sample is calculated for a sub-sample of the full sample that is constituted by those individuals who have completed the nurse visits and with no missing data on all biomeasures used to create allostatic load (8,304 respondents).

[##] The mean of chronological age for the full sample is available only for those who have participated in the nurse visits conducted 5 months on average after the corresponding main waves 2/3 (for direct comparison purposes to the biological age variable that is based on blood-samples collected at nurse visits too).

[###] The mean of biological age for the full sample is calculated for the maximum available sample with valid epigenetic data (3,591 respondents).

Table A.2
Summary statistics: estimation sample versus comparisons sample (without conditioning on biological age).

	Estimation sample[‡]		Comparison sample (without conditioning on biological age)[‡‡]	
	Mean	Obs.	Mean	Obs.
Health care utilisation outcomes				
GP visits (Waves 7–13)	15.590	1,372	16.067	3,818
Outpatient visits (Waves 7–13)	10.083	1,373	9.913	3,827
Any inpatient day (Waves 7–13)[†]	0.302	1,339	0.294	3,705
Any GP visit (Wave 13)[†]	0.580	1,630	0.591	4,587
Any outpatient visit (Wave 13)[†]	0.394	1,632	0.395	4,592
Any inpatient day (Wave 13)[†]	0.065	1,626	0.064	4,573
Predictors				
Initial diagnosed health condition: none[†]	0.642	1,632	0.629	4,594
Initial diagnosed health condition: present[†]	0.358	1,632	0.371	4,594
Allostatic load	-0.491	1,632	-0.618	4,594
Chronological age	52.489	1,632	50.640	4,594
Biological age	44.322	1,632	-	-
Female[†]	0.563	1,632	0.564	4,594
Male[†]	0.437	1,632	0.436	4,594
Log household income	7.382	1,632	7.396	4,594
Secondary/above education[†]	0.830	1,632	0.834	4,594
No/basic education[†]	0.170	1,632	0.166	4,594
Non rented home[†]	0.862	1,632	0.844	4,594
Rent home[†]	0.138	1,632	0.156	4,594

(*continued on next page*)

Table A.2 (*continued*)

	Estimation sample‡		Comparison sample (without conditioning on biological age)‡‡	
	Mean	Obs.	Mean	Obs.
England†	0.820	1,632	0.850	4,594
Wales†	0.086	1,632	0.063	4,594
Scotland†	0.094	1,632	0.087	4,594
Rural†	0.290	1,632	0.269	4,594
Urban†	0.710	1,632	0.731	4,594
Wave 2†	0.618	1,632	0.714	4,594
Wave 3†	0.382	1,632	0.286	4,594

† Dichotomous variable.

‡ Estimation sample as in Table 1; represents the pooled sample of Wave 2 (UKHLS)/ Wave 3 (BHPS) respondents that are followed up at subsequent waves, to obtain future health care use measures, and constrained having non-missing information on the explanatory covariates, on allostatic load, and on biological age. The exact estimation sample (for each health care utilisation regression model) depends on item missingness for the health care utilisation measures.

‡‡ Pooled Wave 2 (UKHLS)/ Wave 3 (BHPS) respondents that are followed up at subsequent waves, to obtain future health care use measures, and constrained having non-missing information on the explanatory covariates and on allostatic load. As opposed to the estimation sample, this sample does not condition on the availability of biological age data.

Table A.3
Selected predictors listed in order of selection: sensitivity analysis using all five epigenetic biological age measures.

GP visits (Waves 7–13)	Outpatient visits (Waves 7–13)	Any inpatient day (Waves 7–13)
Allostatic load	Allostatic load	"PhenoAge"
Initial diagnosed health condition: none	Chronological age	Allostatic load
Male	Initial diagnosed health condition: none	Secondary/above education
Not rented home	Wave 2	Initial diagnosed health condition: none
Secondary/above education	Secondary/above education	"PhenoAge" squared
	Not rented home	
Wave 2	Wales	Log household income
Male*allostatic load	Male	-
"PhenoAge"	Male*Allostatic load	-
-	England	-
-	Log household income	
-	Rural	-

Table A.4
Selected predictors listed in order of selection: sensitivity analysis for GP visits.

Panel A: GP visits (Waves 7–13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Poisson LASSO**
Allostatic load	Allostatic load	Allostatic load	Allostatic load
Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none
Male	Male	Male	Male
Not rented home	Not rented home	Not rented home	Not rented home
Secondary/above education	Secondary/above education	Wave 2	Secondary/above education; Male*Allostatic load
Wave 2; Biological age; Male*Allostatic load	Wave 2; Biological age; Male*Allostatic load	Secondary/above education	Wave 2; Biological age
Scotland	Scotland	Biological age	Scotland
-	-	Male*Allostatic load	-
Panel B: Any GP visit (Wave 13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Probit LASSO**
Allostatic load	Allostatic load	Initial diagnosed health condition: none	Allostatic load
Initial diagnosed health condition: none	Initial diagnosed health condition: none	Allostatic load	Initial diagnosed health condition: none
Male*Allostatic load	Male*Allostatic load	Rural; Biological age	Biological age; Male*Allostatic load
Biological age	Biological age	Male	Rural
Rural	Rural	Male*Allostatic load	Male
Male	Male	Not rented home	Not rented home
Not rented home	Not rented home	England	Log household income
Log household income	Log household income	-	England
England	-	-	-

Table A.5
Selected predictors listed in order of selection: sensitivity analysis for outpatient visits.

Panel A: Outpatient visits (Waves 7–13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Poisson LASSO**
Allostatic load	Allostatic load	Chronological age	Allostatic load
Chronological age	Chronological age	Allostatic load	Chronological age
Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none
Not rented home; Chronological age squared	Not rented home; Chronological age squared	Wave 2	Not rented home
Wave 2; Secondary/above education	Wave 2; Secondary/above education	Not rented home; Chronological age squared	Wave 2
Wales	Wales	Wales	Secondary/above education
Male*Allostatic load	Male*Allostatic load	Log household income	Chonological age squared
Male	Male	-	Male*Allostatic load
Log household income	Log household income	-	Wales
-	England	-	Male
	Rural	-	Log household income
-	-	-	England
-	-	-	Rural
Panel B: Any outpatient visit (Wave 13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Probit LASSO**
Chronological age; Male*Allostatic load	Chronological age; Male*Allostatic load	Chronological age	Chronological age; Male*Allostatic load
Allostatic load	Allostatic load	Male*Allostatic load	Allostatic load
Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none
Male	Male	Male	Male
Rural	Rural	Rural	Rural

Table A.6
Selected predictors listed in order of selection: sensitivity analysis for inpatient days.

Panel A: Any inpatient day (Waves 7–13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Probit LASSO**
Biological age	Biological age	Biological age	Biological age
Allostatic load	Allostatic load	Allostatic load	Allostatic load
Secondary/above education	Secondary/above education	Secondary/above education	Secondary/above education
Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none	Initial diagnosed health condition: none
Biological age squared	Biological age squared	Biological age squared	Log household income
Log household income	Log household income	-	Biological age squared

Panel B: Any inpatient day (Wave 13)			
5 CV folds	**20 CV folds**	**Adaptive LASSO**	**Probit LASSO**
Biological age	Biological age	Biological age	Biological age
Allostatic load	Allostatic load	Allostatic load	Allostatic load
Biological age squared	Biological age squared	Biological age squared	Biological age squared

Table A.7
Selected predictors listed in order of selection: sensitivity analysis when allostatic load is excluded from the set of predictors.

GP visits (Waves 7–13)	**Outpatient visits (Waves 7–13)**	**Any inpatient day (Waves 7–13)**
Initial diagnosed health condition: none	Chronological age	Biological age
Male	Initial diagnosed health condition: none	Secondary/above education
Biological age	Secondary/above education; Not rented home	Initial diagnosed health condition: none
Not rented home	Wave 2	Log household income
Wave 2	Chronological age squared	Biological age squared
Secondary/above education	Male	-
-	Wales	-
-	Log household income	-

References

Ahrens, A., Hansen, C.B., Schaffer, M.E., 2020. lassopack: Model selection and prediction with regularized regression in Stata. Stata J. 20 (1), 176–235.

Bafei, S.E.C., Shen, C., 2023. Biomarkers selection and mathematical modeling in biological age estimation. NPJ. Aging 9 (1), 13.
Bao, Y., Gorrie-Stone, T., Kumari, M., 2022. Understanding Society: Waves 2-3 Nurse Health Assessment 'Epigenetic Clocks' derived from DNA methylation, 2010-2012. University of Essex, U.K. User Guide, Version 1.
Belloni, A., Chernozhukov, V., 2013. Least squares after model selection in high-dimensional sparse models. Bernoulli 521–547.
Benzeval, M., Aguirre, E., Kumari, M., 2023. Understanding society: health, biomarker and genetic data. Fisc. Stud. 44 (4), 399–415.
Brilleman, S.L., Gravelle, H., Hollinghurst, S., Purdy, S., Salisbury, C., Windmeijer, F., 2014. Keep it simple? Predicting primary health care costs with clinical morbidity measures. J. Health Econ. 35, 109–122.
Cameron, A.C., Trivedi, P.K, 2022. Microeconometrics Using Stata, 2nd ed. Stata Press, College Station, TX.
Carreras, M., Ibern, P., Inoriza, J.M., 2018. Ageing and healthcare expenditures: exploring the role of individual health status. Health Econ. 27 (5), 865–876.
Carrieri, V., Jones, A.M., 2017. The income–health relationship 'beyond the mean': new evidence from biomarkers. Health Econ. 26 (7), 937–956.
Chen, B.H., Marioni, R.E., Colicino, E., Peters, M.J., Ward-Caviness, C.K., Tsai, P.C., Roetker, N.S., Just, A.C., Demerath, E.W., Guan, W., Bressler, J., 2016. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany. NY) 8 (9), 1844.
Costa-Font, J., Vilaplana-Prieto, C., 2020. More than one red herring'? Heterogeneous effects of ageing on health care utilisation. Health Econ. 29, 8–29.
Crimmins, E., Kim, J.K., Fisher, J., Faul, J.D., 2020. HRS Epigenetic Clocks – Release 1. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, U.S.A.
Cui, L., Xie, X., Shen, Z., Lu, R., Wang, H., 2018. Prediction of the healthcare resource utilisation using multi-output regression models. IISE Trans. Healthc. Syst. Eng. 8 (4), 291–302.
Dalgaard, C.J., Strulik, H., 2014. Optimal aging and death: understanding the Preston curve. J. Eur. Econ. Assoc. 12 (3), 672–701.
Davillas, A., Jones, A.M., Benzeval, M., 2019. The income-health gradient: evidence from self-reported health and biomarkers in understanding society. Panel Data Econometrics. Academic Press, pp. 709–741.
Davillas, A., Pudney, S., 2020a. Using biomarkers to predict healthcare costs: evidence from a UK household panel. J. Health Econ. 73, 102356.
Davillas, A., Pudney, S., 2020b. Biomarkers, disability and health care demand. Econ. Human Biol. 39, 100929.
De Meijer, C., Koopmanschap, M., d'Uva, T.B., Van Doorslaer, E., 2011. Determinants of long-term care spending: age, time to death or disability? J. Health Econ. 30 (2), 425–438.
Demirer, I., Schmidt, B., Schramm, S., Erbel, R., Jöckel, K.H., Pförtner, T.K., 2021. Does allostatic load predict incidental coronary events differently among sexes? Compr. Psychoneuroendocrinol. 8, 100089.
di Martino, E., Honey, S., Bradley, S.H., Ali, O.M., Neal, R.D., Scott, S.E., 2023. Understanding GPs' referral decisions for younger patients with symptoms of cancer: a qualitative interview study. Br. J. General Pract.
Dupras, C., Beck, S., Rothstein, M.A., Berner, A., Saulnier, K.M., Pinkesz, M., Joly, Y., 2019. Potential (mis) use of epigenetic age estimators by private companies and public agencies: human rights law should provide ethical guidance. Environ. Epigenet. 5 (3), dvz018.
Faul, J.D., Kim, J.K., Levine, M.E., Thyagarajan, B., Weir, D.R., Crimmins, E.M., 2023. Epigenetic-based age acceleration in a representative sample of older Americans: associations with aging-related morbidity and mortality. Proc. National Acad. Sci. 120 (9), e2215840120.
Fransquet, P.D., Wrigglesworth, J., Woods, R.L., Ernst, M.E., Ryan, J., 2019. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. Clin. Epigenetics. 11, 1–17.
Grossman, M., 1972. The Demand for Health: a Theoretical and Empirical Investigation. National Bureau of Economic Research, Columbia University Press.
Hannum, G., Guinney, J., Zhao, L., Zhang, L.I., Hughes, G., Sadda, S., Zhang, K., 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell 49 (2), 359–367.
Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical learning with sparsity. Monographs Stat. Appl. Probability 143 (143), 8.
Henderson, R.H., French, D., Maughan, T., Adams, R., Allemani, C., Minicozzi, P., Lawler, M., 2021. The economic burden of colorectal cancer across Europe: a population-based cost-of-illness study. Lancet Gastroenterol. Hepatol. 6 (9), 709–722.
Horvath, S., 2013. DNA methylation age of human tissues and cell types. Genome Biol. 14, 1–20.
Horvath, S., Raj, K., 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nature Rev. Genet. 19 (6), 371–384.
Howdon, D., Rice, N., 2018. Health care expenditures, age, proximity to death and morbidity: Implications for an ageing population. J. Health Econ. 57, 60–74.
Jain, P., Binder, A.M., Chen, B., Parada, H., Gallo, L.C., Alcaraz, J., LaCroix, A.Z., 2022. Analysis of epigenetic age acceleration and healthy longevity among older US women. JAMa Netw. Open. 5 (7), e2223285.
Jylhävä, J., Pedersen, N.L., Hägg, S., 2017. Biological Age Predictors, 21. EBioMedicine, pp. 29–36.
Levine, M.E., Lu, A.T., Quach, A., Chen, B.H., Assimes, T.L., Bandinelli, S., Horvath, S., 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging 10 (4), 573.
Li, A., Koch, Z., Ideker, T., 2022. Epigenetic aging: Biological age prediction and informing a mechanistic theory of aging. J. Intern. Med. 292 (5), 733–744.
Lin, Q., Weidner, C.I., Costa, I.G., Marioni, R.E., Ferreira, M.R., Deary, I.J., Wagner, W., 2016. DNA methylation levels at individual age-associated CpG sites can be indicative for life expectancy. Aging 8 (2), 394.
Marx, G., Koens, S., von Dem Knesebeck, O., Scherer, M., 2022. Age and gender differences in diagnostic decision-making of early heart failure: results of a mixed-methods interview-study using video vignettes. BMJ Open. 12 (3), e054025.
McCrory, C., Fiorito, G., McLoughlin, S., Polidoro, S., Cheallaigh, C.N., Bourke, N., Karisola, P., Alenius, H., Vineis, P., Layte, R., Kenny, R.A., 2020. Epigenetic clocks and allostatic load reveal potential sex-specific drivers of biological aging. J. Gerontol.: Series A 75 (3), 495–503.
OECD, 2021. Demographic Trends, in Health at a Glance 2021: OECD Indicators. OECD Publishing, Paris. https://doi.org/10.1787/9989e95c-en.
Ohlsson, C., Labrie, F., Barrett-Connor, E., Karlsson, M.K., Ljunggren, O., Vandenput, L., Tivesten, A., 2010. Low serum levels of dehydroepiandrosterone sulfate predict all-cause and cardiovascular mortality in elderly Swedish men. J. Clin. Endocrinol. Metabolism 95 (9), 4406–4414.
Padula, W.V., Kreif, N., Vanness, D.J., Adamson, B., Rueda, J.D., Felizzi, F., Crown, W., 2022. Machine learning methods in health economics and outcomes research—the PALISADE checklist: a good practices report of an ISPOR task force. Value in Health 25 (7), 1063–1080.
Parker, Haley W., Abreu, Alyssa M., Sullivan, Mary C., Vadiveloo, Maya K., 2022. Allostatic load and mortality: a systematic review and meta-analysis. Am. J. Prev. Med. 63, 131–140.
Phyo, A.Z.Z., Espinoza, S.E., Murray, A.M., Fransquet, P.D., Wrigglesworth, J., Woods, R.L., Ryan, J., 2024. Epigenetic age acceleration and the risk of frailty, and persistent activities of daily living (ADL) disability. Age Ageing 53 (6), afae127.
Ryan, C.P., 2021. Epigenetic clocks": Theory and applications in human biology. Am. J. Human Biol. 33 (3), e23488.
Shepherd, D., 2017. Capitation based funding of general practice is not fit for purpose. BMJ 358.
Shorrocks, A.F., 2013. Decomposition procedures for distributional analysis: a unified framework based on the sahpley value. J. Econ. Inequal. 11, 99–126.
Sibley, L.M., Glazier, R.H., 2012. Evaluation of the equity of age–sex adjusted primary care capitation payments in Ontario. Canada. Health Policy 104 (2), 186–192.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. Royal Stat. Society Series B: Stat. Methodol. 58 (1), 267–288.
Van Doorslaer, E., Koolman, X., Jones, A.M., 2004. Explaining income-related inequalities in doctor utilisation in Europe. Health Econ. 13 (7), 629–647.
Wang, W., Dearman, A., Bao, Y., Kumari, M., 2023. Partnership status and positive DNA methylation age acceleration across the adult lifespan in the UK. SSM-Population Health 24, 101551.
Zavala, D.V., Dzikowski, N., Gopalan, S., Harrington, K.D., Pasquini, G., Mogle, J., Scott, S.B., 2024. Epigenetic age acceleration and chronological age: associations with cognitive performance in daily life. J. Gerontol.: Series A 79 (1), glad242.
Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 101 (476), 1418–1429.
Zweifel, P., Felder, S., Meiers, M., 1999. Ageing of population and health care expenditure: a red herring? Health Econ. 8 (6), 485–496.