

<https://doi.org/10.1038/s44324-025-00078-x>

Metabolomic-based aging clocks

Check for updates

A. Ibáñez de Opakua¹, R. Conde², A. de Diego², M. Bizkarguenaga², N. Embade², S. C. Lu³, J. M. Mato^{2,4} ✉ & O. Millet^{1,2,4} ✉

Molecular aging clocks estimate biological age from molecular biomarkers and often outperform chronological age in predicting health outcomes. Types include epigenetic, transcriptomic, proteomic, and metabolomic clocks. NMR-based metabolomic clocks provide a non-invasive, high-throughput platform to assess metabolic health. We summarize key NMR-based models and present a new approach that combines high predictive accuracy with clinical interpretability, identifying disease-specific metabolic distortions and supporting risk stratification and early detection of accelerated aging.

Development and validation of molecular aging clocks

A molecular aging clock is a biomarker-based computational model designed to estimate an organism's biological age—a measure of physiological state that often predicts healthspan, disease risk, and mortality more accurately than chronological age (the number of years since birth). Biological age reflects the cumulative impact of both intrinsic (genetic) and extrinsic (environmental/lifestyle) factors, as well as the progressive functional decline of cells and organ systems due to processes like cellular senescence, mitochondrial dysfunction, and loss of proteostasis. These clocks are generated by analyzing molecular changes that accumulate over time in diverse biological samples (e.g., blood, saliva, and tissues) collected from individuals of different ages. Machine learning and statistical models are then trained to correlate these molecular patterns with chronological age (first-generation clocks) or to predict health-related outcomes (second-generation clocks).

Main types of molecular aging clocks

Epigenetic clocks

DNA methylation patterns have been identified as some of the most reliable aging biomarkers. Examples include Horvath and Hannum et al. clocks^{1,2}, and newer models like McCrory et al.³, which refine predictions based on different tissue types and populations. Epigenetic clocks tend to be very accurate in estimating chronological age; they often have limited ability to identify the key biological processes controlling ageing⁴. Intriguingly, it has been suggested that epigenetic-based aging clocks may be sensitive to stochastic variation accumulation in the DNA⁵.

Transcriptomic clocks

Changes in gene expression are also strong indicators of aging. Clocks based on transcriptomic data, such as Singh et al. and Lu et al.^{6,7}, capture shifts in regulatory networks driving functional decline.

Remarkably, this method was able to predict accelerated aging in a small cohort of progeria patients⁸.

Proteomic clocks

Aging is associated with variations in protein abundance and post-translational modifications. Examples include Argentieri's clock, which used proteomics to predict age-related functional status, multimorbidity, and mortality risk across geographically and genetically diverse populations⁹. Proteomic-based aging clocks have been recently reviewed¹⁰.

Metabolomic clocks

Metabolite level changes provide insights into metabolic alterations associated with aging. For example, Bucaciuc and coworkers developed a metabolomic clock that detects aging-induced shifts such as NAD⁺ depletion and mitochondrial dysfunction¹¹.

Additional markers under investigation include telomere length, senescence-associated secretory phenotype factors, chromatin remodeling, and microbiome shifts, among others^{12,13}.

Validation and application of aging clocks

Once generated, molecular aging clocks undergo validation on independent cohorts to ensure their accuracy and generalizability¹⁴. Their primary applications include: (1) First-generation clocks: Estimate chronological age based on molecular biomarkers. A significant discrepancy between biological and chronological age may indicate accelerated or decelerated aging, which can signal health risks or resilience. (2) Second-generation clocks: Predict health-related factors such as mortality risk, frailty, disease susceptibility, response to therapies, and the effects of lifestyle interventions (e.g., diet, exercise, and pharmacological treatments like metformin or rapamycin). As molecular aging clocks advance, the aim is to develop personalized multi-omic models that integrate genomic, epigenomic,

¹ATLAS Molecular Pharma, Parque Tecnológico de Bizkaia, Ed. 800, 48160 Derio, Spain. ²Precision Medicine and Metabolism Laboratory, CIC bioGUNE, Basque Research and Technology Alliance, Parque Tecnológico de Bizkaia, Ed. 800, 48160 Derio, Spain. ³Karsh Division of Gastroenterology and Hepatology, Cedars-Sinai Medical Center, Los Angeles, CA, USA. ⁴CIBERehd, Madrid, 28025, Spain. ✉ e-mail: director@cicbiogune.es; omillet@cicbiogune.es

transcriptomic, proteomic, and metabolomic data for a comprehensive assessment of aging and longevity¹⁵.

Limitations of single biological age measurements

Despite their potential, molecular aging clocks have several limitations that impact their accuracy, applicability, and interpretation¹⁶:

Correlation vs. causation

Many aging clocks rely on molecular markers that correlate with age, but it remains unclear whether these markers actively drive aging or are merely byproducts of the aging process. For instance, DNA methylation changes are strongly associated with aging but may simply reflect age-related physiological shifts rather than directly influencing lifespan. In this context, it is important to define the influence of genetic variation and environmental factors on changes that are used to build-up the aging clock model.

Population-specific bias

Aging clocks trained on one population may not generalize well to others due to differences in genetics, lifestyle, and environmental exposures. Factors such as ethnicity, diet, environment, and socio-economic status can influence aging trajectories, potentially leading to biased predictions when applying a model across diverse groups. Retraining or recalibrating aging clocks with population-specific data is often necessary, and data integration from large and diverse longitudinal population studies is highly desirable.

Uncertain impact of longevity interventions

While reductions in biological age following lifestyle interventions (e.g., diet, exercise, caloric restriction, or pharmacological treatments like metformin or rapamycin) suggest potential anti-aging effects, it is unclear whether these changes translate into actual lifespan extension or improved long-term health.

Limitations of single-time measurements

A single biological age measurement provides only a snapshot of aging at a given moment and has limited value in tracking aging dynamics. Since aging is a continuous process, transient influences such as illness, stress, or recent lifestyle changes can cause short-term fluctuations in biological age estimates, leading to potential misinterpretations.

Since aging is dynamic, a single measurement cannot capture the rate of aging or its changes over time. Repeated measurements are necessary to assess whether aging is accelerating, decelerating, or progressing normally. Advanced models, such as DunedinPACE¹⁷, estimate the pace of aging by tracking individual decline over decades. These approaches offer deeper insights into biological aging, beyond a simple age estimate.

NMR-based metabolomic aging clocks

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical tool widely employed in metabolomics for the qualitative and quantitative analysis of small-molecule metabolites in biological samples, including plasma, urine, cerebrospinal fluid, and tissue extracts. Its non-destructive nature, high reproducibility, and minimal sample preparation requirements make it particularly suited for comprehensive metabolic profiling. While the technique has intermediate sensitivity compared to other platforms, it allows for reliable quantification of metabolite concentrations. Serum and plasma are of particular interest due to their tightly regulated homeostasis and clinically relevant concentration ranges. Metabolic profiling of serum is especially informative for assessing central metabolism and glycemic control, key processes linked to aging and longevity¹⁸. Moreover, NMR offers unprecedented resolution in lipoprotein subclass characterization, and the combination of metabolites and lipoproteins as determined by NMR spectroscopy provide insights into cardiometabolic risk and metabolic syndrome¹⁹. In addition, specific serum markers measurable by NMR, such as GlycA, GlycB, and SPC, serve as robust inflammatory biomarkers. Collectively, NMR spectroscopy represents a powerful platform for

constructing molecular models of metabolic age and developing metabolomics-based aging clocks.

NMR spectroscopy has been employed to develop various models of biological aging. Hertel and colleagues²⁰ introduced a metabolomic approach to estimate biological age (termed the metabolic age score), based on urine samples analyzed using ¹H NMR spectroscopy. Their study included 4068 urine samples collected from Caucasian donors across Central Europe. While the model achieved a linear correlation between chronological age and metabolic age, its predictive accuracy was limited, likely due to a regression-to-the-mean effect. Furthermore, urine is not an ideal biofluid for long-term tracking of metabolic age, as its composition is highly sensitive to short-term influences such as diet and medication.

An NMR-based analysis of serum and plasma samples from the Estonian cohort (17,345 individuals), which included mortality data over a five-year period, revealed a non-linear association between chronological age and a biomarker score that could predict mortality independently of its cause (cardiovascular, nonvascular, or cancer-related deaths)²¹. Similarly, in a study using the FINRRISK cohort (44,168 individuals), a combination of 14 metabolites was identified as predictive biomarkers of mortality²². In turn, Dimitri and coworkers analyzed a small cohort of healthy individuals and people with Parkinson's disease in order to investigate the metabolic aging associated with the neurodegenerative disorder²³.

Another study²⁴ utilized over 18,000 serum samples collected from 26 Dutch hospitals to develop a metabolomics-based age predictor, termed *metaboAge*, aimed at estimating an individual's biological age. The model was built on 56 metabolic features and included a diverse sample set comprising both healthy donors and patients, as defined by the inclusion criteria. As a result, the model reflects deviations from typical population norms, which is consistent with its relatively modest correlation with chronological age (with a Pearson coefficient of 0.65). In line with these findings, Ala-Korpela and colleagues analyzed serum samples from two distinct Finnish cohorts to investigate the existence of slow and accelerated aging regions; however, they were unable to clearly identify such regions²⁵.

Finally, metabolic age models have been developed based on the analysis of metabolic variables measured in blood samples from a large number of individuals across nine cohort studies in the UK and Finland, with participants ranging in age from 24 to 86 years^{26,27}. The same studies also conducted a meta-analysis of existing models and found that while metabolic age models are only moderately correlated with chronological age in independent populations, they offer additional predictive power for morbidity and mortality beyond that of chronological age alone.

Towards an NMR-based global metabolomic health test

At CIC bioGUNE, we initiated a precision medicine program involving a cohort of 13,500 individuals from the Basque Country (AKRIBEA cohort)²⁸. This untargeted study recruits participants exclusively based on their employment within the Mondragón Corporation (<https://www.mondragon-corporation.com/en/>), a federation of worker cooperatives headquartered in the Basque region of Spain, which employs approximately 31,000 individuals locally. Recruitment and biological sample collection are conducted during routine annual medical examinations. Mondragón Corporation operates across four sectors—finance, industry, retail, and knowledge—thereby enabling the collection of samples from a diverse range of occupational environments.

This cohort served as the foundation for developing an NMR-based metabolomic aging clock model, which was further supplemented with additional samples to ensure balanced representation across the full age spectrum, yielding a final dataset encompassing approximately 20,000 individuals. For model construction, we utilized one-dimensional ¹H-NMR (NOESY) spectra and applied a robust ensemble stacking machine learning approach to predict chronological age (see Materials and methods). This strategy effectively mitigated the common issue of regression to the mean, thereby enhancing generalizability (Fig. 1). The current version of the model achieves a Pearson correlation coefficient exceeding 0.90 between metabolic

Fig. 1 | Correlation between chronological age and metabolic age for the test set of the 1D ¹H NOESY-based model. (A, blue) The cross-validation sets of the model based on quantified metabolites and clinical parameters (B, red), and the corresponding independent test set (C, green). The NOESY-based model achieved a Pearson correlation coefficient of 0.92 on the test set, with more than 75% of individuals exhibiting prediction errors smaller than 10 years. Slightly lower performance was observed for the metabolite-based model in both cross-validation ($R = 0.87$) and test evaluations ($R = 0.88$).

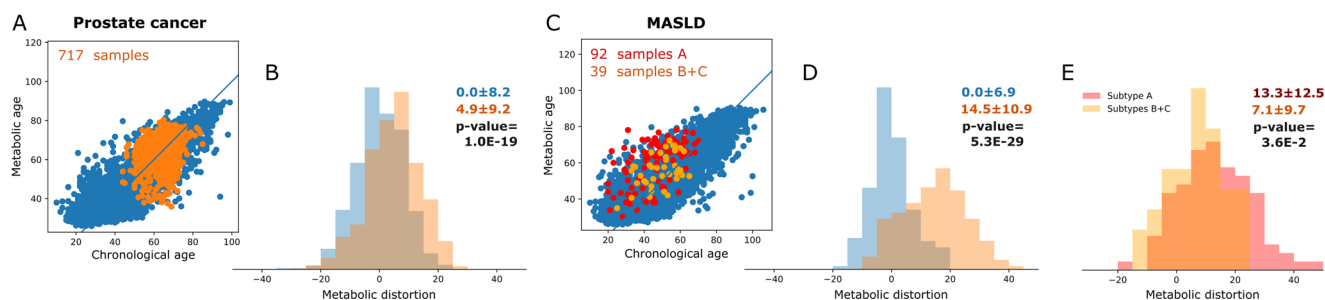
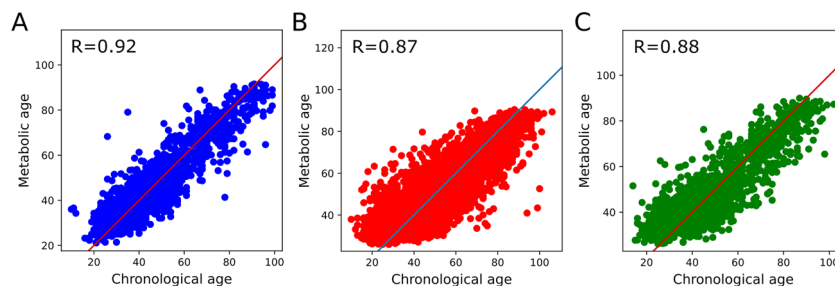


Fig. 2 | Metabolic age distributions in individuals with prostate cancer and metabolic dysfunction-associated steatotic liver disease (MASLD). Samples from individuals with prostate cancer are overlaid on those from a male reference population in (A) (orange and blue, respectively), and histograms of the differences between metabolic and chronological age (metabolic distortion) for these groups are shown in (B). C MASLD samples overlaid on a reference population, with subtype A in red and subtypes B + C in orange, while D presents histograms of metabolic

distortion for these MASLD subtypes alongside a matched reference population (orange and blue, respectively). Differences in metabolic distortion between MASLD subtype A and subtypes B + C are shown in (E). Mean and standard deviation values of metabolic distortion for individuals with prostate cancer and MASLD (combined subtypes A + B + C) are indicated in orange; corresponding values for their respective reference populations are indicated in blue. p -Values from Kolmogorov–Smirnov tests assessing distribution differences are shown in black.

and chronological age, with significantly reduced prediction error compared to our previous versions of the model (Fig. 1A). A slightly less accurate but more interpretable version of the model, based on selected metabolites and clinical parameters extracted from the NOESY spectra, reaches a Pearson correlation just below 0.90 (Fig. 1B, C).

Importantly, individuals with various pathologies, including chronic metabolic and oncological conditions, exhibited characteristic deviations from their predicted metabolic age. In prostate cancer, a disease more likely to develop in older men²⁹, 717 cases were analyzed, ranging from 55 to 75 years of age, with a mean chronological age of approximately 67 years (Fig. 2A). The clinical and biochemical characteristics of the patients with prostate cancer included in the study have previously been described³⁰. In this cohort, the metabolic distortion histogram, representing the differences between metabolic and chronological ages, revealed a significant shift toward older metabolic ages ($+4.9 \pm 9.2$ years, p -Value = 1.0×10^{-19}), while maintaining a comparable distribution width (Fig. 2B), suggesting an overall acceleration of metabolic aging in prostate cancer patients.

In contrast, in metabolic dysfunction-associated steatotic liver disease (MASLD), a condition affecting both men and women across a wide age range³¹, the analyzed cases ($N = 169$, of which 131 had a defined subtype) ranged from 20 to 70 years of age, with a mean chronological age of approximately 65 years (Fig. 2C). The clinical and biochemical characteristics of the patients with MASLD included in this study have previously been described³². The corresponding metabolic distortion histogram (Fig. 2D) displayed a broader distribution, indicating greater heterogeneity in metabolic distortion and a significant shift toward older metabolic ages in MASLD ($+14.5 \pm 10.9$ years, p -value = 1.9×10^{-29}). These findings are consistent with our previous work identifying distinct serum lipidomic profiles among MASLD patients, independent of histological disease severity. Specifically, we classified patients into three metabolic subtypes (A–C), and found that individuals with MASLD with subtype A exhibited

lower serum very low-density lipoprotein levels and cardiovascular disease risk than those with subtypes B and C³².

Accordingly, in Fig. 2C, MASLD subtype A samples are shown in red, and subtypes B + C in orange. A comparison of their metabolic distortion histograms (Fig. 2E) revealed a modest but statistically significant difference: $+13.3 \pm 12.5$ years for subtype A vs. $+7.1 \pm 9.7$ years for subtype B + C (p -value = 3.6×10^{-2}). These results suggest differences in metabolic aging among MASLD subtypes that are not captured by standard clinical evaluations, including histological assessment and conventional serum biomarkers. The primary limitation of this finding is the relatively small number of MASLD individuals with available subtype classification. These distinct patterns underscore the model's potential as a complementary marker for uncovering clinically relevant but otherwise undetected disease heterogeneity.

The ultimate objective is to bring clinical relevance to the metabolic profiling space. To enhance interpretability and extend the analytical scope beyond age prediction, we developed a metabolite quantification pipeline based on fast-acquisition two-dimensional J-resolved NMR spectra. This method enabled the identification and quantification of up to 49 serum metabolites by minimizing signal overlap, a common limitation of one-dimensional spectra. Additionally, the NOESY spectrum inherently contains valuable information on lipoprotein subclasses and inflammatory markers (*vide supra*). Building on this, we employed supervised ensemble learning models trained on the same spectral data to estimate 25 additional clinical parameters, encompassing both directly measurable biomarkers (e.g., albumin and CRP) and indirectly inferred indices (e.g., calcium levels, glomerular filtration rate, leukocyte count). To better understand model behavior and clinical drivers of the age prediction, we used Shapley Additive exPlanations (SHAP), a method that quantifies the average contribution of each feature to the model's predictions) to interpret the feature importance of the stacking ensemble model (Fig. 3). This summary plot highlights the top 10 most influential features ranked by their average absolute SHAP

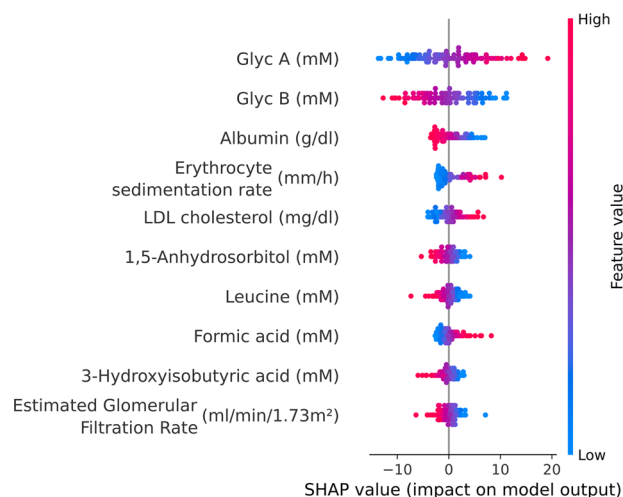


Fig. 3 | Feature importance based on SHAP values for the stacking ensemble machine learning model trained on 49 quantified metabolites and 25 predicted clinical parameters. Features are ranked from most to least influential based on their average absolute SHAP values. Each point represents a SHAP value for an individual observation, with color indicating the original feature value (red = high, blue = low). A positive SHAP value indicates that the feature contributes to a higher predicted age, while a negative SHAP value indicates a contribution to a lower predicted age. This plot illustrates both the magnitude and direction of each feature's impact on model predictions.

values. For example, high values of Glyc A are associated with higher predicted age, while low levels of albumin also lead to higher age predictions, evident by the inversion in color gradients. Both markers are linked to inflammation: Glyc A is a direct inflammatory biomarker, and albumin levels tend to decrease during inflammatory responses, highlighting the relevance of inflammation as a key physiological process in aging. While this integration improves clinical interpretability, it does come with a modest trade-off in predictive performance, with the model's *R* value decreasing from 0.92 to 0.88 (Fig. 1B, C). This hybrid dataset ultimately serves as a powerful platform for integrating metabolic and clinical data, substantially enhancing the interpretability and translational potential of the model (Fig. 3).

³³Previous studies using NMR-based data for age prediction have shown limited performance, with maximum reported Pearson correlation coefficients around 0.78^{20,24,26}. In contrast, our models substantially exceed this threshold, achieving correlations of 0.92 (NOESY-based) and 0.88 (metabolite-based), highlighting significant methodological advancements. Nevertheless, recent studies^{33,34} argue that the value of an aging clock may lie less in its ability to accurately predict chronological age and more in how strongly its age delta (i.e., the difference between predicted and chronological age) relates to clinically relevant outcomes. On the other hand, our findings resonate with those of Zhang et al.³³, particularly in the importance assigned to GlycA, albumin, and lipoprotein-related features.

Our best-performing model (a stacking ensemble that integrates a Ridge linear regression with an ExtraTreesRegressor) bears a structural resemblance to the Cubist model used by Mutz et al.³⁴, which combines decision tree-derived rules with localized linear models. Both approaches leverage the strengths of tree-based methods for capturing complex, non-linear interactions, while incorporating linear regression to model additive effects and enhance interpretability.

While NMR-derived metabolic models demonstrate strong performance in predicting chronological age, several limitations should be acknowledged. On one hand, the models are always developed and validated on specific cohorts, and their generalizability to more diverse populations, including different ethnicities, geographic regions, or health conditions, is required. Second, the interpretability of these models is inherently limited compared to simpler, hypothesis-driven approaches. Future work should

focus on external validation, exploration of causative biological mechanisms, and potential clinical utility in longitudinal or interventional settings.

In conclusion, the integration of high-throughput NMR spectroscopy with advanced machine learning techniques has enabled the development of a robust, interpretable, and clinically meaningful metabolic age model (Biogune's model). By combining predictive NMR-based metabolomic aging clocks with detailed metabolite profiling, lipoprotein subclass analysis, and estimations of standard clinical parameters, this approach moves beyond basic diagnostics toward a comprehensive health assessment platform. The incorporation of direct and inferred biomarkers bridges the gap between metabolic phenotyping and clinical utility, offering a promising avenue for early disease detection, personalized health monitoring, and risk stratification. As this initiative continues to evolve, its scalability, cost-effectiveness, and non-invasive nature position it as a compelling candidate for implementation in routine clinical practice, ultimately advancing the goals of precision medicine.

Materials and methods

Cohorts

The study leveraged several cohorts to ensure robust coverage across a wide age spectrum, including healthy individuals and patients with relevant conditions. Participants were from Southern Europe (Portugal, Spain, and Italy).

- **AKRIBEA:** A cohort of individuals recruited from the Mondragón Corporation, a federation of worker cooperatives in the Basque Country (Spain), covering diverse sectors (finance, industry, retail, and knowledge). Participants were recruited during routine annual medical check-ups. This cohort serves as a representative sample of the working-age general population. Sample size: 13,545
- **DDM-Madrid:** Women aged 39–50 from the Madrid region (Spain), who attended gynecological screenings at the Madrid Salud Diagnostic Center between June 2013 and May 2015. Participants were invited via telephone. Sample size: 933
- **Biosilver:** Individuals recruited from geriatric centers, health centers, and hospitals across Spain as part of an aging-focused study. Fasting serum samples were collected between 2023 and 2025. Sample size: 411
- **Liver-Bible:** Healthy blood donors from Milan (Italy), recruited for a comprehensive screening of liver, metabolic, and cardiovascular health. Sample size: 1641
- **SPBB:** Samples selected from biobanks of Spain integrated into the ISCIII Biobanks and Biomodels Platform to supplement under-represented older age groups. Sample size: 3353
- **AGEPORTUGAL:** Participants from an aging study conducted in Portuguese geriatric centers. Sample size: 247
- **Prostate Cancer:** Individuals with prostate cancer recruited at Basurto University Hospital (Spain), with samples managed by the Basque Biobank for Research (BIOEF). Sample size: 717
- **LITMUS:** Patients with biopsied metabolic dysfunction-associated steatotic liver disease (MASLD) from clinical centers in the UK and Italy. Sample size: 169

To protect patient confidentiality, all data were double-coded prior to analysis.

Blood collection and serum preparation

Venous blood was collected from fasting participants. Serum samples were processed according to standardized operating protocols (Bizkarguenaga et al.,²⁸) and stored at -80°C until analysis. Briefly, blood was allowed to clot at room temperature, centrifuged, and the supernatant serum was aliquoted into cryovials for long-term storage.

NMR spectroscopy

¹H-NMR spectra were acquired using Bruker Avance III HD and Neo IVDx 600 MHz spectrometers, equipped with BBI probes and SampleJet™ automation, maintaining a sample temperature of 5°C . Calibration and quality control followed the procedures described by³⁵, ensuring spectral

reproducibility and quantitative reliability. For each sample, a standard one-dimensional NOESY spectrum (with solvent presaturation) and a fast-acquisition two-dimensional J-resolved spectrum were recorded.

Metabolite and clinical parameter determination

To enhance interpretability, two complementary quantification strategies were employed:

1. Metabolite quantification: Concentrations of 49 metabolites were estimated based on the integration of distinct peaks in the 2D J-resolved NMR spectra, which offer improved spectral resolution and reduced peak overlap compared to 1D spectra. Quantifications were validated through spiking experiments using serum samples. However, since these quantifications are used as input for machine learning models, the absolute accuracy of the reported values is less critical than their internal precision and consistency across samples. The quantified metabolites include:

1,5-Anhydrosorbitol, 2-Aminobutyric acid, 2-Hydroxybutyric acid, 2-Oxoglutaric acid, 3-Hydroxybutyric acid, 3-Hydroxyisobutyric acid, Acetic acid, Acetoacetic acid, Acetone, Alanine, Arginine, Asparagine, Aspartate, Betaine, Choline, Citric acid, Creatine, Creatinine, Cystine, D-Galactose, Dimethylamine, Dimethylsulfone, Ethanol, Formic acid, Glucose, Glutamic acid, Glutamine, Glycerol, Glycine, Histidine, Isoleucine, Lactic acid, Leucine, Lysine, Methanol, Methionine, Myo-inositol, N,N-Dimethylglycine, Ornithine, Phenylalanine, Proline, Pyruvic acid, Sarcosine, Serine, Succinic acid, Threonine, Trimethylamine-N-oxide, Tyrosine, Valine.

2. Clinical parameter estimation: For 25 clinical parameters, machine learning models were trained using 1D NMR spectra as input and clinically measured values as targets. Models were optimized using TPOT (Tree-based Pipeline Optimization Tool), a genetic programming framework for automated machine learning (autoML), which selects and tunes pipelines for optimal predictive performance. The performance of each model is reported as the Pearson correlation coefficient (R) between the predicted and measured values, along with the number of samples used for training (indicated in parentheses). The estimated parameters and their respective performance are as follows:

Albumin (R = 0.94, n = 557), Apolipoprotein B (R = 0.82, n = 532), Bilirubin (R = 0.70, n = 579), Calcium (R = 0.81, n = 556), C-reactive protein (R = 0.80, n = 1565), Erythrocyte sedimentation rate (R = 0.78, n = 22,148), Erythrocytes (R = 0.73, n = 22,113), Estimated glomerular filtration rate (R = 0.79, n = 1394), Fructosamine (R = 0.77, n = 212), Glyc A (R = 0.99, n = 14,288), Glyc B (R = 0.98, n = 14,288), HDL cholesterol (R = 0.97, n = 25,984), Hemoglobin (R = 0.80, n = 23,724), Iron (R = 0.79, n = 511), LDL cholesterol (R = 0.96, n = 22,672), Leukocytes (R = 0.62, n = 22,671), Lipoprotein(a) (R = 0.95, n = 3526), Platelets (R = 0.62, n = 23,725), SPC (R = 1.00, n = 14,288), Total cholesterol (R = 0.97, n = 25,781), Total protein (R = 0.79, n = 1285), Transferrin (R = 0.90, n = 511), Triglycerides (R = 0.95, n = 26,054), Urate (R = 0.79, n = 22,262), Urea (R = 0.99, n = 557).

Data analysis

To reduce age-related sampling bias, five age ranges were defined: 0–30, 31–40, 41–50, 51–65, and 66+. Random selection was used to equalize the number of samples per group, yielding a final balanced dataset of 9500 samples (1900 per age group).

Two parallel machine learning models were developed to predict chronological age:

- Model 1: Utilized the 1D 1H-NMR spectra (NOESY) as input.
- Model 2: Used the quantified metabolite concentrations and clinical parameters.

Both models followed a common pipeline: data standardization followed by a stacking ensemble method, combining a Ridge linear regression model with an ExtraTreesRegressor (a tree-based ensemble model). Model 1

prioritized predictive performance, while Model 2 was designed for interpretability.

Model optimization was performed using a cross-validation strategy. First, 20% of the total dataset was set aside as an independent test set. The remaining 80% was used for model training and internal evaluation through fivefold cross-validation.

The results of the five cross-validation folds for Model 2 are presented together in Fig. 1B, while the final performance on the held-out test set is shown in Fig. 1A, C for Models 1 and 2, respectively.

Interpretability was achieved using SHAP (SHapley Additive exPlanations) values, which quantify the contribution of each feature to individual predictions. However, caution must be exercised when interpreting SHAP values in biological data due to potential multicollinearity, as biological variables are often highly correlated, which can distort feature importance estimates.

Data availability

Data will be available upon request to the corresponding authors.

Abbreviations

NMR	nuclear magnetic resonance spectroscopy
NOESY	nuclear Overhauser effect spectroscopy
SHAP	SHapley Additive exPlanations

Received: 16 May 2025; Accepted: 20 July 2025;

Published online: 03 September 2025

References

1. Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
2. Horvath, H. & Horvath, S. DNA Methylation Age of Human Tissues and Cell Types. *Genome Biol.* <http://genomebiology.com/2013/14/10/R115> (2013).
3. McCrory, C. et al. GrimAge outperforms other epigenetic clocks in the prediction of age-related clinical phenotypes and all-cause mortality. *J. Gerontol.* **76**, 741–749 (2021).
4. Prosz, A. et al. Biologically informed deep learning for explainable epigenetic clocks. *Sci. Rep.* **14**, 1306 (2024).
5. Meyer, D. H. & Schumacher, B. Aging clocks based on accumulating stochastic variation. *Nat. Aging* **4**, 871–885 (2024).
6. Singh, S. P. et al. Machine learning based classification of cells into chronological stages using single-cell transcriptomics. *Sci. Rep.* **8**, 17156 (2018).
7. Lu, J. et al. Heterogeneity and transcriptome changes of human CD8+ T cells across nine decades of life. *Nat. Commun.* **13**, 5128 (2022).
8. Fleischer, J. G. et al. Predicting age from the transcriptome of human dermal fibroblasts. *Genome Biol.* **19**, 221 (2018).
9. Argentieri, M. A. et al. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat. Med.* **30**, 2450–2460 (2024).
10. Johnson, A. A., Shokhirev, M. N., Wyss-Coray, T. & Lehallier, B. Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age. *Ageing Res. Rev.* <https://doi.org/10.1016/j.arr.2020.101070> (2020).
11. Bucaciuc Mracica, T. et al. MetaboAge DB: a repository of known ageing-related changes in the human metabolome. *Biogerontology* **21**, 763–771 (2020).
12. Palmer, R. D. Aging clocks & mortality timers, methylation, glycomic, telomeric and more. A window to measuring biological age. *Ageing Med.* <https://doi.org/10.1002/agm2.12197> (2022).
13. Lin, M., Egli, C., Dulai, A. S. & Sivamani, R. K. Critical review of aging clocks and factors that may influence the pace of aging. *Front. Aging* <https://doi.org/10.3389/fragi.2024.1487260> (2024).

14. Liang, R., Tang, Q., Chen, J. & Zhu, L. Epigenetic Clocks: Beyond Biological Age, Using the Past to Predict the Present and Future. *Aging Dis* <https://doi.org/10.14336/AD.2024.1495>. (2024)
15. Mavromatis, L. A. et al. Multi-omic underpinnings of epigenetic aging and human longevity. *Nat. Commun.* **14**, 2236 (2023).
16. Bell, C. G. et al. DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* **20**, 249 (2019).
17. Belsky, D. W. et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* **11**, e73420 (2022).
18. López-Otín, C., Galluzzi, L., Freije, J. M. P., Madeo, F. & Kroemer, G. Metabolic control of longevity. *Cell* **166**, 802–821 (2016).
19. Gil-Redondo, R. et al. MetSCORE: a molecular metric to evaluate the risk of metabolic syndrome based on serum NMR metabolomics. *Cardiovasc. Diabetol.* **23**, 1–13 (2024).
20. Hertel, J. et al. Measuring biological age via metabolomics: the metabolic age score. *J. Proteome Res.* **15**, 400–410 (2016).
21. Fischer, K. et al. Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 Persons. *PLoS Med.* **11**, e1001606 (2014).
22. Deelen, J. et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. *Nat. Commun.* **10**, 3346 (2019).
23. Dimitri, G. M., Meoni, G., Tenori, L., Luchinat, C. & Lió, P. On behalf of the PROPAG-AGEING Consortium NMR spectroscopy combined with machine learning approaches for age prediction in healthy and Parkinsons disease cohorts through metabolomic fingerprints. *Appl. Sci* **12**, 8954 (2022).
24. Van Den Akker, E. B. et al. Metabolic age based on the BBMRI-NL 1H-NMR metabolomics repository as biomarker of age-related disease. *Circ. Genom. Precis. Med.* <https://doi.org/10.1161/CIRCGEN.119.002610> (2020).
25. Ala-Korpela, M. et al. Cross-sectionally calculated metabolic aging does not relate to longitudinal metabolic changes—support for stratified aging models. *J. Clin. Endocrinol. Metab.* **108**, 2099–2104 (2023).
26. Lau, C. H. E. et al. NMR metabolomic modeling of age and lifespan: a multicohort analysis. *Aging Cell* **23**, e14164 (2024).
27. Shang, X. et al. Metabolomic age and risk of 50 chronic diseases in community-dwelling adults: a prospective cohort study. *Aging Cell* **23**, e14125 (2024).
28. Bizkarguenaga, M. et al. Prospective metabolomic studies in precision medicine: the AKRIBEA Project. In *Handbook of Experimental Pharmacology* vol. 277 pp 275–297 (2022).
29. Sekhoacha, M. et al. Prostate cancer review: genetics, diagnosis, treatment options, and alternative approaches. *Molecules* **27**, 5730 (2022).
30. Bruzzzone, C. et al. 1H-NMR-based urine metabolomics reveals signs of enhanced carbon and nitrogen recycling in prostate cancer. *J. Proteome Res.* **19**, 2419–2428 (2020).
31. Younossi, Z. M. et al. Global consensus recommendations for metabolic dysfunction-associated steatotic liver disease and steatohepatitis. *Gastroenterology* <https://doi.org/10.1053/j.gastro.2025.02.044> (2025).
32. Martínez-Arranz, I. et al. Metabolic subtypes of patients with NAFLD exhibit distinctive cardiovascular risk profiles. *Hepatology* **76**, 1121–1134 (2022).
33. Zhang, S. et al. A metabolomic profile of biological aging in 250,341 individuals from the UK Biobank. *Nat. Commun.* **15**, 8081 (2024).
34. Mutz, J., Iniesta, R. & Lewis, C. M. Metabolomic age (MileAge) predicts health and life span: a comparison of multiple machine learning algorithms. *Sci. Adv.* **10**, eadp3743 (2024).
35. Dona, A.C. et al. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal. Chem.* **86**, 9887–9894 (2014).

Acknowledgements

Support was provided by the Department of Economic Development and Infrastructures of the Government of the Autonomous Community of the Basque Country (Elkartek, bg2021, bg2023) and the Plan Complementario en biotecnología aplicada en Salud. This work was supported by NIH grant R01DK123763 (O. Millet, S.C. Lu, J.M. Mato), the Grant CEX2021-001136-S funded by MICIU/AEI/10.13039/501100011033, the project PID2023-146132OB-I00 funded by MICIU/AEI/10.13039/501100011033 (J.M. Mato) and by ERDF, EU and Ciberehd_ISCIII_MINECO, funded by the Instituto de Salud Carlos III (J.M. Mato).

Author contributions

R.C., M.B., Ad.D. and N.E. collected the experimental datasets, A.I.O. analyzed the data and A.I., S.L., J.M.M., and O.M. wrote the paper. All authors revised and approved the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J. M. Mato or O. Millet.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025