Multimodal contrastive regression for organ-resolved biological age prediction

Veronika Ecker[1,2] VERONIKA.ECKER@ISS.UNI-STUTTGART.DE
[1] *Institute of Signal Processing and System Theory, University of Stuttgart*
[2] *Diagnostic and Interventional Radiology, University Hospital of Tuebingen*

Tim Förster[1] ST189453@STUD.UNI-STUTTGART.DE
Sergios Gatidis[2] SGATIDIS@STANFORD.EDU
Thomas Küstner[*2] THOMAS.KUESTNER@MED.UNI-TUEBINGEN.DE
Bin Yang[*1] BIN.YANG@ISS.UNI-STUTTGART.DE

Editors: Under Review for MIDL 2026

Abstract

Aging is complex and heterogeneous. A person's chronological age (CA) cannot reflect individual aging trajectories which prompted investigations for biological age (BA). Existing BA models often rely on single organs or limited biomarkers, restricting holistic assessment. Overcoming these constraints requires multimodal models that better reflect the distributed nature of aging. Accurate estimation of BA therefore requires models that capture the combined influence of structural, functional, and physiological factors on the aging process. BA progresses heterogeneously across organs, motivating models that integrate multiple data types to capture this complexity. We propose a multimodal contrastive regression framework that jointly learns from MRI and structured clinical variables to estimate organ-specific BA. A contrastive regression loss structures the latent space to reflect continuous age differences both within and across modalities. Applied to a large population cohort, this approach produces well-structured latent embeddings, improves age estimation relative to unimodal systems, and reveals characteristic age-gap patterns in healthy and diseased subgroups. The results demonstrate that combining MRI with complementary tabular features strengthens BA estimation and supports a comprehensive multi-organ view of aging.

Keywords: Multimodal Data, Contrastive Regression, Biological Age, Multi-organ, Magnetic Resonance Imaging

1. Introduction

The risk of developing many prevalent diseases, including cardiovascular disease (Lakatta, 2002; North and Sinclair, 2012), cancer (DePinho, 2000), and neurodegenerative disorders (Katzman and Saitoh, 1991; Corrada et al., 1995), increases significantly with advancing age. Consequently, accurate assessment of an individual's aging status provides prognostically and diagnostically relevant information, offering better insights into health status and disease risk. Conventionally, chronological age is used as the proxy for aging, but it only reflects time since birth and does not capture the individual biological changes that

---

[*] Contributed equally

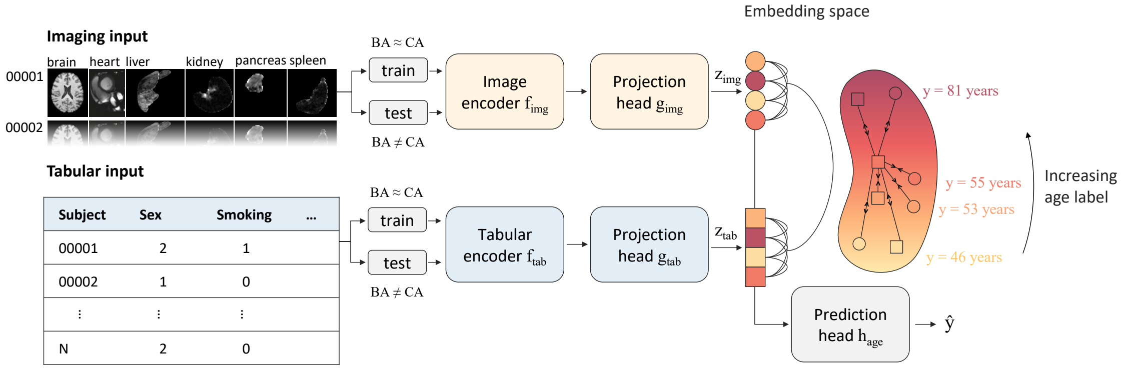
drive functional decline. These individual differences result from a complex interplay of genetic, environmental, and lifestyle factors. Biological age reflects the functional state of organs and tissues beyond chronological age, offering insights into individual health, vulnerability to disease, and deviations from typical aging trajectories. The development of a standardized, robust method to determine biological age remains an unresolved challenge in medicine. One complicating factor is that different organs age at different rates (Tian et al., 2023), requiring the determination of an organ-specific biological age to fully characterize an individual's health trajectory.

Prior efforts to estimate biological age have relied primarily on molecular profiles (Horvath, 2013; Hannum et al., 2013; Meyer and Schumacher, 2021), laboratory measurements (Le Goallec and Patel, 2019; Sagers et al., 2020), or functional tests (Rahman and Adjeroh, 2019; Nakamura et al., 1990). These methods have contributed important insights, but each captures only part of the aging process and often produces one global estimate rather than organ-specific values. The recent rise of deep learning has made it possible to derive aging information directly from the rich, high-dimensional content of medical images, such as Magnetic Resonance Imaging (MRI) (Ecker et al., 2024; Smith et al., 2019), Computed Tomography (CT) (Azarfar et al., 2024), and X-ray (Ieki et al., 2022). Given the multifactorial nature of the aging process, it is hypothesized that leveraging multiple sources of information is critical for accurate prediction. If accompanying electronic health records or tabular information exists, there is a significant opportunity to jointly leverage multimodal data to enhance the biological age estimation task. This concept has also been proven beneficial in other domains (Boehm et al., 2022; Mohsen et al., 2022).

Previous multimodal learning studies have explored pairing structured clinical information with imaging inputs including MRI (Hager et al., 2023), CT (Fang et al., 2021), or chest radiographs (Zhou et al., 2023). Other works link tabular variables with time series signals like electrocardiographic data (Stym-Popper et al., 2025). These approaches vary primarily in the encoder design and in the fusion strategy. Most applications focus on classification tasks, either to detect disease, such as Alzheimer's disease (Chen et al., 2020a; K. P. and P., 2021) and myocardial infarction (Sharma et al., 2021), or to predict disease onset, including breast cancer metastasis (Ma et al., 2020), chronic hepatitis (Yanru Bai et al., 2016), Alzheimer's disease (Bhagwat et al., 2018), or mortality risk (Aljouie et al., 2021).

In multimodal learning, self-supervised contrastive learning has emerged as a powerful method for pre-training medical models (Hager et al., 2023). The key idea is that embeddings organize input data in latent space so that semantically similar samples stay closer together, effectively grouping subjects with similar properties relevant for downstream tasks. While this naturally supports classification, regression tasks benefit from additional structure to reflect the ordering of continuous target variables. To address this, Dufumier et al. (2021) introduced a contrastive regression loss, using chronological age as a continuous proxy label for classification tasks. This strategy was later adopted by Barbano et al. (2023) to estimate brain age from MRI data.

Building on these advances, we introduce a novel adaptation of contrastive regression to estimate organ-specific biological age that integrates multimodal data. Our approach combines MRI scans with complementary tabular features, including lifestyle factors, blood-based measures, and imaging-derived biomarkers, to predict biological age in six organ systems: brain, heart, liver, kidneys, spleen, and pancreas. By jointly exploiting imaging and tabular



Subject	Sex	Smoking	...
00001	2	1	
00002	1	0	
⋮	⋮	⋮	
N	2	0	

Figure 1: Schematic of the contrastive framework for organ-specific biological age (BA) estimation. Organ-specific imaging and tabular inputs are split into a healthy training cohort and a mixed (healthy and diseased) test cohort. Contrastive pre-training orders embeddings by age distance. After pre-training, a prediction head is trained using the ordered feature embeddings as input to predict the age $\hat{y}$. During training we use healthy controls where we can assume that BA ≈ chronological age (CA).

data, our results demonstrate that multimodal integration improves age prediction performance over unimodal approaches, highlighting the value of combining complementary sources of information for a more accurate and organ-specific assessment of biological aging. Training on a healthy UK Biobank cohort with chronological age allows the model to capture normative aging patterns, while testing is performed on a mixed subgroup.

2. Materials and Methods

2.1. Data

The study uses the UK Biobank cohort, a large resource of health information from adults aged 44 to 83 years across the UK. In addition to extensive data on environmental exposures, lifestyle, medical history, biological samples, and physical assessments, the project also offers a substantial imaging dataset. The available scans include 3D MPRAGE brain MRI, 2D+time cardiac cine MRI, and 3D abdominal Dixon VIBE MRI from roughly 70,000 participants. Image pre-processing included skull stripping and alignment to MNI152 space for brain scans (Evans et al., 1993). After segmentation, cardiac and abdominal images were cropped to extract organ specific volumes of the heart, liver, spleen, kidneys, and pancreas. Abdominal organs were additionally processed with background masking to prevent shortcut learning through fat tissue outside the target structures. Brain and cardiac scans were processed as single channel inputs, while abdominal images used four channels for the Dixon contrasts. For organ-specific tabular data, we selected features linked to organ aging such as imaging-derived biomarkers as well as blood-based measurements and lifestyle factors. Continuous variables were z-score normalized, and cardinal categorical variables were onehot encoded. For downstream analysis, disease age for several common conditions was derived from ICD-10 codes and corresponding diagnosis dates in electronic health records. Diagnoses were categorized as pre-existing or future depending on whether they occurred before or after the imaging time point used for biological age estimation.

2.2. Multimodal integration

Our framework adopts the basic SimCLR architecture (Hager et al., 2023; Chen et al., 2020b) as a backbone, consisting of separate encoders and projection heads for image and tabular data (Figure 1). For each subject, a pair of inputs, imaging data x_{img} and tabular data x_{tab}, is processed through their respective encoder and projection head branches. The image encoder $f_{\text{img}}(\cdot)$ is based on a ResNet18 architecture (He et al., 2016) modified for 3D data by replacing standard 2D convolutional layers with 3D convolutions. This adaptation allows the model to leverage spatial context in volumetric data, such as brain or abdominal scans, or spatiotemporal context in 2D+time sequences, like cardiac imaging. The tabular encoder $f_{\text{tab}}(\cdot)$ is implemented as a multilayer perceptron with five dense layers, using batch normalization and ReLU activations. The resulting embeddings are then mapped into a shared projection space through projection heads g_{img} and g_{tab}. Each projection head consists of two dense layers with layer normalization and LeakyReLU activation, yielding final projections $z_{\text{img}} \in \mathbb{R}^{128}$ and $z_{\text{tab}} \in \mathbb{R}^{128}$:

$$\begin{aligned} z_{\text{img}} &= g_{\text{img}}(f_{\text{img}}(x_{\text{img}})) \\ z_{\text{tab}} &= g_{\text{tab}}(f_{\text{tab}}(x_{\text{tab}})) \end{aligned} \tag{1}$$

These projections are then optimized in a shared embedding space to align the modalities according to their age label difference.

2.3. Multimodal contrastive regression loss

Conventional contrastive representation learning for classification defines positive pairs when samples originate from the same subject or class and negative pairs otherwise. The model is then optimized to pull embeddings of positive pairs closer together and push negative pairs farther apart in the embedding space. However, in regression tasks, the labels are continuous, requiring the embedding space to follow a continuous structure rather than a categorical scheme. Instead of distinct positive and negative pairs, the degree to which a pair should be pulled together or pushed apart should gradually scale based on the closeness of their label distance. To adapt the standard InfoNCE loss (Oord et al., 2018) for this setting, we use a weighting factor based on a Gaussian kernel function $w_{i,j} = K(y_i - y_j)$ (Dufumier et al., 2021; Barbano et al., 2023), where y_i and y_j are the continuous labels of samples i and j, respectively. In a batch of size B, all possible pairs are successively considered. Every sample j relative to the anchor i is initially treated as a potential positive, but the weighting factor reduces the influence of pairs with large label distance. Only samples with sufficiently similar age labels to y_i contribute meaningfully to the loss. This weighting scheme transforms the standard InfoNCE-based loss into a contrastive regression loss that enforces a continuous structure on the latent representation space. The tolerance margin for label similarity is controlled by the standard deviation σ of the Gaussian kernel, which should be selected in relation to the batch size as well as the range and distribution of the age labels in the dataset. To extend this formulation to the multimodal setting, the loss is applied across modalities and within each modality. The aim is to align paired imaging and tabular representations while also ordering samples within each modality according to their ground truth labels y_i. For an anchor i from modality $m \in \{1, 2, ..., M\}$, candidates from all modalities are considered, leading to loss term l_{i_m} of sample i and modality m:

$$l_{i_m} = \sum_{n=1}^{M} \sum_{j=1}^{B} \mathbb{1}_{j \neq i, n \neq m} \frac{w_{i,j}}{\sum_{j=1}^{B} \mathbb{1}_{j \neq i} w_{i,j}} \left(- \ln \frac{s_{i_m, j_n}}{\sum_{n=1}^{M} \sum_{j=1}^{B} \mathbb{1}_{j \neq i, n \neq m} s_{i_m, j_n}} \right), \tag{2}$$

where the similarity is defined as $s_{i_m, j_n} = \exp(sim(z_{i_m}, z_{j_n}))$ using the cosine similarity $sim(\cdot, \cdot)$. In our experimental setup, we consider two modalities and thus set $M = 2$, but the loss generalizes naturally to settings involving a single modality or more than two modalities. The final multimodal contrastive regression loss is obtained by summing over all imaging and tabular anchors:

$$\mathcal{L} = \frac{1}{BM} \sum_{i=1}^{B} \sum_{m=1}^{M} l_{i_m} \tag{3}$$

2.4. Biological age prediction

Following contrastive pre-training, the learned ordered representations z_m, $m \in \{\text{img}, \text{tab}\}$, are concatenated and serve as input to the age prediction head (see Figure 1). Contrary to the standard SimCLR approach of removing projection heads before prediction, we keep them. The projections z_m preserve better age ordering than the raw embeddings, providing a more informative input for the regression task. Therefore, the encoder and projection heads could, in principle, be merged. The age prediction head h_{age} is implemented as a multilayer perceptron with four dense layers, each followed by batch normalization and LeakyReLU activations, and a final linear layer to output the predicted age $\hat{y}_i$. The model is trained using mean squared error (MSE) loss with respect to the chronological age label.

$$\mathcal{L} = \frac{1}{B} \sum_{i=1}^{B} (y_i - \hat{y}_i)^2 \tag{4}$$

Since biological age lacks direct ground truth, we approximate it using chronological age within a carefully selected organ-specific healthy cohort, assuming biological age $\approx$ chronological age for this population. When the model is applied to a broader test set including both healthy and diseased individuals, the age difference between them can be interpreted as accelerated or decelerated biological aging.

3. Experimental Setting

Separate models were trained for each organ. A healthy subcohort was selected by excluding participants with diagnoses listed in predefined, organ-specific disease categories derived from electronic medical records within the UK Biobank. The resulting healthy subset was divided into train and test partitions in a 70/30 split. For the final analysis of predicted biological age, the healthy test samples were merged with the disease-affected groups that

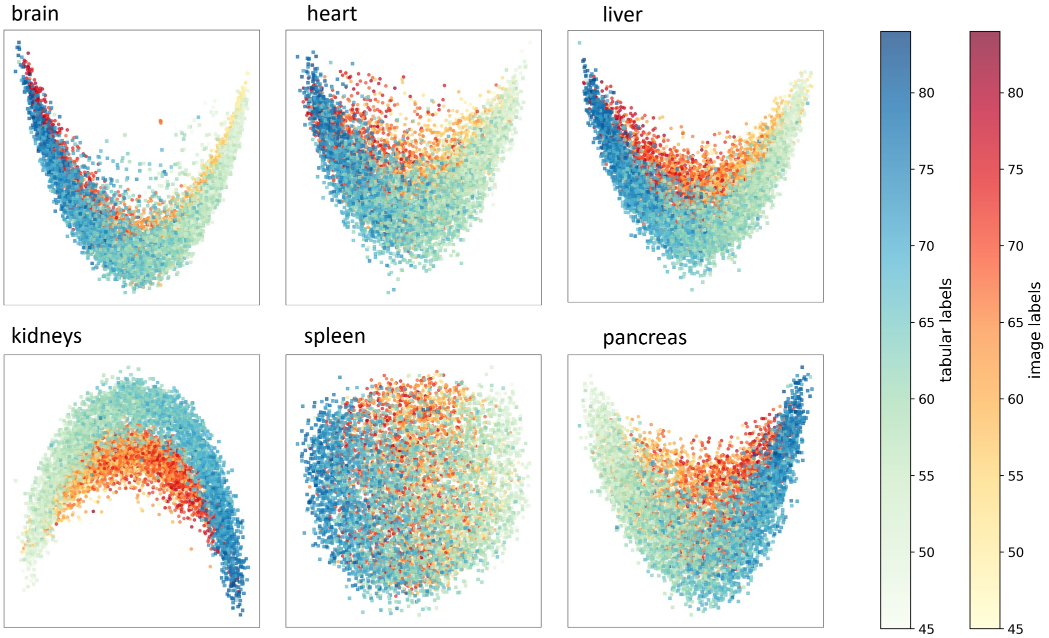


Figure 2: Visualization of global patterns in imaging and tabular feature embeddings after contrastive pre-training for all organs using principal component analysis plots. Color grading shows age labels for imaging (yellow–red) and tabular (green–blue) features.

were previously excluded. Because imaging availability and exclusion rules varied by organ, the number of training cases ranged from 19,128 for the brain to 30,641 for the spleen. To reduce overfitting, augmentation strategies were applied, including resizing, cropping, mirroring, and gamma adjustments. Contrastive pre-training was conducted for 12 epochs with a batch size of 100 on two A100 GPUs, using the Adam optimizer (Kingma and Ba, 2014) and a learning rate schedule, followed by 10 epochs of training of the age prediction head. To assess the role of multimodal fusion, unimodal models ($M = 1$) using only imaging or only tabular features were also trained under the same settings as multimodal training. For a fair comparison and to gain further insights into tabular performance, the tabular features in the unimodal experiments were divided into imaging-derived biomarkers and non-imaging variables, including lifestyle information and blood-based measurements. After training the age predictions were bias-corrected using linear regression to account for the inherent tendency of regression models for extreme values to shift towards the mean.

4. Results and Discussion

Pre-training was effective across all organs, as shown by PCA visualizations that project the high dimensional embeddings into a lower dimensional space while preserving broad structural patterns (Figure 2). The loss function encourages both imaging and tabular representations to follow the chronological age order, ensuring consistency within each modality

Table 1: Mean absolute errors of predicted ages in healthy the test set for multimodal approach and unimodal ablations.

	Brain	**Heart**	**Liver**	**Kidneys**	**Spleen**	**Pancreas**
Multimodal	**2.60**	**3.66**	**3.01**	**2.80**	**3.73**	**3.33**
Imaging	2.65	4.49	4.50	5.49	6.00	5.45
Image Biomarkers	3.64	5.41	5.47	5.45	5.46	5.50
Non-imaging Biomarkers	4.69	4.15	3.40	3.56	3.86	3.44

as well as between modalities. Tabular encoders showed robust performance across organs. Imaging encoders were more sensitive to how easily aging information can be extracted. In the brain, strong markers of aging, including volumetric reductions in gray and white matter (Kaye et al., 1992), contributed to a well structured embedding space with few outliers. In comparison, organs like the spleen and pancreas showed more outliers, since their visible age-related changes can be mainly attributed to an overall volume loss.

After training, performance was assessed on an healthy subset of the test cohort. Including diseased subjects could be misleading, as lower mean absolute errors in mixed cohorts do not necessarily reflect better prediction. Mean absolute errors (MAE) for multimodal models and the three unimodal ablations are summarized in Table 1. Imaging and image-biomarker unimodal models worked well in organs where aging produces clear structural changes, such as the brain, but were less effective in organs with very limited age-related imaging signals, such as the kidney, spleen, and pancreas. In organs with strong imaging markers of aging, full images supported better predictions. In organs with weak signals, including the spleen and kidney, handcrafted image biomarkers performed better than full image despite the limited number of abdominal image biomarkers (see Supplementary Figure 6). Tabular models without imaging data were consistent across organs and generally outperformed imaging models, with the exception of the brain where imaging information is particularly informative. Multimodal training effectively leveraged complementary information in all organs, surpassing all unimodal performances in all organs.

By including participants with diseases, predicted age can be interpreted as an indicator of biological age. The predicted age gap, defined as the difference between predicted and chronological age, provides insight into accelerated or decelerated aging. When comparing age gaps between patients with common diseases and healthy controls, average predicted age gaps appeared elevated in certain diseased subgroups (Figure 3). This pattern was specifically examined for Alzheimer's disease in the brain, type II diabetes in the pancreas, and chronic kidney disease in the kidneys. Higher predicted age gaps were observed both in participants with pre-existing conditions and in those who developed the disease after data collection, indicating potential prognostic relevance.

Beyond disease specific effects, aging patterns can also be examined across organs to understand systemic relationships (Figure 4). Brain aging occurs relatively independently, whereas abdominal organs and the heart are moderately correlated. Strong correlations between the left and right kidneys reflect expected behavior of bilateral organs. Overall, organs age heterogeneously but shared patterns are visible across the system.

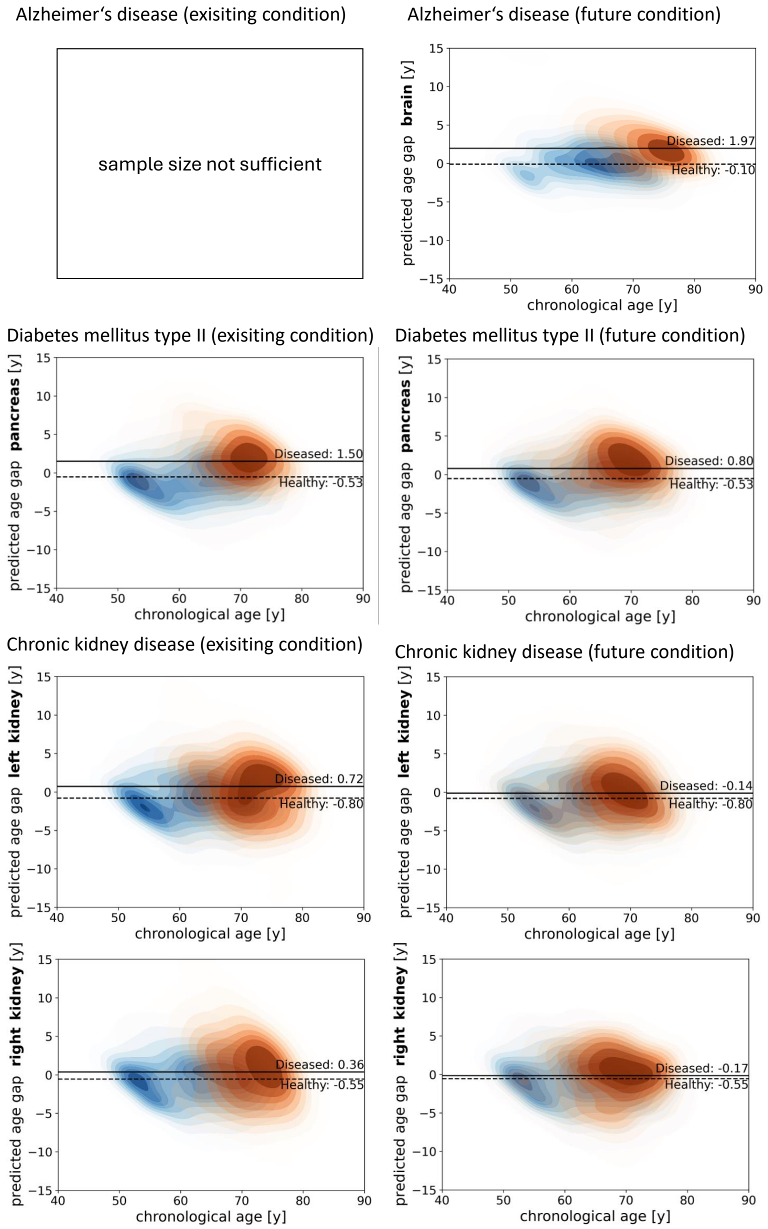


Figure 3: Kernel density estimations (KDE) of predicted age gaps (= predicted age - chronological age) across organs for specific health outcomes. KDE plots show predicted age gaps in test subjects with pre-existing or future diagnoses (orange) overlaid to healthy controls (blue). Mean values are shown for the healthy cohort (dashed line) and diseased cohort (solid line).

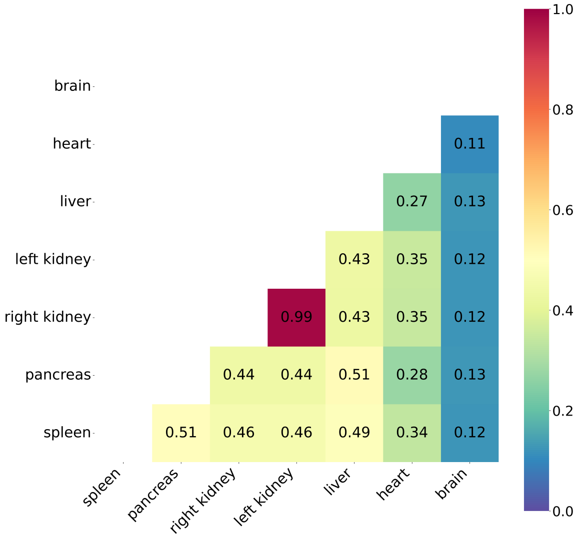


Figure 4: Heatmap of Pearson correlation coefficients between predicted age gaps of brain, heart, left kidney, right kidney, pancreas, and spleen. All results are statistically significant with $p < 2.38 \times 10^{-3}$ (21 tests, Bonferroni-corrected).

5. Conclusion

In summary, our study demonstrates that combining MRI and tabular data provides a powerful framework for estimating organ-specific biological age. By aligning multimodal features through contrastive regression, the models capture complementary aging information across six organ systems: brain, heart, liver, kidneys, spleen, and pancreas. Evaluations reveal that multimodal models generally outperform unimodal approaches, identify accelerated aging in diseased participants compared to healthy individuals, and uncover correlated aging patterns across organs. Performance is occasionally limited when images provide little age-related information, causing the image encoder to impair the overall model. Advanced fusion techniques may overcome this limitation and strengthen the predictive capabilities of multimodal approaches.

Acknowledgments

This work was carried out under UK Biobank Application 60520. We thank all participants who took part in the UKBB study and the staff in this research program. The work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project #428219130 and supported under Germany's Excellence Strategy (EXC 2064/1, #390727645).

References

Abdulrhman Fahad Aljouie, Ahmed Almazroa, Yahya Bokhari, Mohammed Alawad, Ebrahim Mahmoud, Eman Alawad, Ali Alsehawi, Mamoon Rashid, Lamya Alomair, Shahad Almozaai, Bedoor Albesher, Hassan Alomaish, Rayyan Daghistani, Naif Khalaf Alharbi, Manal Alaamery, Mohammad Bosaeed, and Hesham Alshaalan. Early prediction of COVID-19 ventilation requirement and mortality from routinely collected baseline chest radiographs, laboratory, and clinical data with machine learning. 14:2017–2033, 2021. ISSN 1178-2390. doi: 10.2147/JMDH.S322431.

Ghazal Azarfar, Seok-Bum Ko, Scott J. Adams, and Paul S. Babyn. Deep learning-based age estimation from chest CT scans. 19(1):119–127, 2024. ISSN 1861-6429. doi: 10.1007/s11548-023-02989-w. URL `https://doi.org/10.1007/s11548-023-02989-w`.

C. A. Barbano, B. Dufumier, E. Duchesnay, M. Grangetto, and P. Gori. Contrastive learning for regression in multi-site brain age prediction. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pages 1–4, 2023. ISBN 1945-8452. doi: 10.1109/ISBI53787.2023.10230733. Journal Abbreviation: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI).

Nikhil Bhagwat, Joseph D. Viviano, Aristotle N. Voineskos, M. Mallar Chakravarty, and Alzheimer's Disease Neuroimaging Initiative. Modeling and prediction of clinical symptom trajectories in alzheimer's disease using longitudinal data. 14(9):e1006376, 2018. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1006376. URL `https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006376`. Publisher: Public Library of Science.

Kevin M. Boehm, Pegah Khosravi, Rami Vanguri, Jianjiong Gao, and Sohrab P. Shah. Harnessing multimodal data integration to advance precision oncology. 22(2):114–126, 2022. ISSN 1474-1768. doi: 10.1038/s41568-021-00408-3.

D. Chen, L. Zhang, and C. Ma. A multimodal diagnosis predictive model of alzheimer's disease with few-shot learning. In *2020 International Conference on Public Health and Data Science (ICPHDS)*, pages 273–277, 2020a. Journal Abbreviation: 2020 International Conference on Public Health and Data Science (ICPHDS).

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020b.

Maria Corrada, Ronald Brookmeyer, and Claudia Kawas. Sources of variability in prevalence rates of alzheimer's disease. 24(5):1000–1005, 1995. ISSN 0300-5771. doi: 10.1093/ije/24.5.1000. URL https://doi.org/10.1093/ije/24.5.1000. _eprint: https://academic.oup.com/ije/article-pdf/24/5/1000/2079392/24-5-1000.pdf.

Ronald A. DePinho. The age of cancer. 408(6809):248–254, 2000. ISSN 1476-4687. doi: 10.1038/35041694. URL https://doi.org/10.1038/35041694.

Benoit Dufumier, Pietro Gori, Julie Victor, Antoine Grigis, Michele Wessa, Paolo Brambilla, Pauline Favre, Mircea Polosan, Colm McDonald, Camille Marie Piguet, Mary Phillips, Lisa Eyler, and Edouard Duchesnay. Contrastive learning with continuous proxy metadata for 3d MRI classification. In Marleen de Bruijne, Philippe C. Cattin, Stéphane Cotin, Nicolas Padoy, Stefanie Speidel, Yefeng Zheng, and Caroline Essert, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pages 58–68. Springer International Publishing, 2021. ISBN 978-3-030-87196-3.

Veronika Ecker, Marcel Früh, Bin Yang, Sergios Gatidis, and Thomas Küstner. Deep regression for biological age estimation in multiple organs: Investigations on 40,000 subjects of the UK biobank. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2255–2259, 2024. ISBN 2379-190X. doi: 10.1109/ICASSP48485.2024.10447990. Journal Abbreviation: ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

A.C. Evans, D.L. Collins, S.R. Mills, E.D. Brown, R.L. Kelly, and T.M. Peters. 3d statistical neuroanatomical models from 305 MRI volumes. In *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817 vol.3, 1993. doi: 10.1109/NSSMIC.1993.373602. URL https://ieeexplore.ieee.org/document/373602.

Cong Fang, Song Bai, Qianlan Chen, Yu Zhou, Liming Xia, Lixin Qin, Shi Gong, Xudong Xie, Chunhua Zhou, Dandan Tu, Changzheng Zhang, Xiaowu Liu, Weiwei Chen, Xiang Bai, and Philip H.S. Torr. Deep learning for predicting COVID-19 malignant progression. 72:102096, 2021. ISSN 1361-8415. doi: 10.1016/j.media.2021.102096. URL https://www.sciencedirect.com/science/article/pii/S1361841521001420.

Paul Hager, J. M. Menten, J., and Daniel Rueckert. Best of both worlds: Multimodal contrastive learning with tabular and imaging data. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 23924–23935, 2023. ISBN 2575-7075. doi: 10.1109/CVPR52729.2023.02291. Journal Abbreviation: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Gregory Hannum, Justin Guinney, Ling Zhao, Li Zhang, Guy Hughes, SriniVas Sadda, Brandy Klotzle, Marina Bibikova, Jian-Bing Fan, Yuan Gao, Rob Deconde, Menzies Chen, Indika Rajapakse, Stephen Friend, Trey Ideker, and Kang Zhang. Genome-wide methylation profiles reveal quantitative views of human aging rates. 49(2):359–367, 2013. ISSN 1097-2765. doi: 10.1016/j.molcel.2012.10.016. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3780611/.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. doi: 10.1109/CVPR.2016.90. URL https://ieeexplore.ieee.org/document/7780459. ISSN: 1063-6919.

Steve Horvath. DNA methylation age of human tissues and cell types. 14(10):3156, 2013. ISSN 1474-760X. doi: 10.1186/gb-2013-14-10-r115. URL https://doi.org/10.1186/gb-2013-14-10-r115.

Hirotaka Ieki, Kaoru Ito, Mike Saji, Rei Kawakami, Yuji Nagatomo, Kaori Takada, Toshiya Kariyasu, Haruhiko Machida, Satoshi Koyama, Hiroki Yoshida, Ryo Kurosawa, Hiroshi Matsunaga, Kazuo Miyazawa, Kouichi Ozaki, Yoshihiro Onouchi, Susumu Katsushika, Ryo Matsuoka, Hiroki Shinohara, Toshihiro Yamaguchi, Satoshi Kodera, Yasutomi Higashikuni, Katsuhito Fujiu, Hiroshi Akazawa, Nobuo Iguchi, Mitsuaki Isobe, Tsutomu Yoshikawa, and Issei Komuro. Deep learning-based age estimation from chest x-rays indicates cardiovascular prognosis. 2(1):1–12, 2022. ISSN 2730-664X. doi: 10.1038/s43856-022-00220-6. URL https://www.nature.com/articles/s43856-022-00220-6. Publisher: Nature Publishing Group.

Muhammed Niyas K. P. and Thiyagarajan P. Alzheimer's classification using dynamic ensemble of classifiers selection algorithms: A performance analysis. 68:102729, 2021. ISSN 1746-8094. doi: 10.1016/j.bspc.2021.102729. URL https://www.sciencedirect.com/science/article/pii/S1746809421003268.

Robert Katzman and Tsunao Saitoh. Advances in alzheimer's disease. 5(3):278–286, 1991. ISSN 0892-6638. doi: 10.1096/fasebj.5.3.2001787. URL https://doi.org/10.1096/fasebj.5.3.2001787. Publisher: John Wiley & Sons, Ltd.

J. A. Kaye, C. DeCarli, J. S. Luxenberg, and S. I. Rapoport. The significance of age-related enlargement of the cerebral ventricles in healthy men and women measured by quantitative computed x-ray tomography. 40(3):225–231, 1992. ISSN 0002-8614. doi: 10.1111/j.1532-5415.1992.tb02073.x.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. 2014.

Edward G. Lakatta. Age-associated cardiovascular changes in health: Impact on cardiovascular disease in older persons. 7(1):29–49, 2002. ISSN 1573-7322. doi: 10.1023/A:1013797722156. URL https://doi.org/10.1023/A:1013797722156.

Alan Le Goallec and Chirag J. Patel. Age-dependent co-dependency structure of biomarkers in the general population of the united states. 11(5):1404–1426, 2019. ISSN 1945-4589. doi: 10.18632/aging.101842.

Wenjuan Ma, Xin Wang, Guijun Xu, Zheng Liu, Zhuming Yin, Yao Xu, Haixiao Wu, Vladimir P. Baklaushev, Karl Peltzer, Henian Sun, Natalia V. Kharchenko, Lisha Qi, Min Mao, Yanbo Li, Peifang Liu, Vladimir P. Chekhonin, and Chao Zhang. Distant metastasis prediction via a multi-feature fusion model in breast cancer. 12(18):18151–18162, 2020. ISSN 1945-4589. doi: 10.18632/aging.103630.

David H. Meyer and Björn Schumacher. BiT age: A transcriptome-based aging clock near the theoretical limit of accuracy. 20(3):e13320, 2021. ISSN 1474-9726. doi: 10.1111/acel.13320.

Farida Mohsen, Hazrat Ali, Nady El Hajj, and Zubair Shah. Artificial intelligence-based methods for fusion of electronic health records and imaging data. 12(1):17981, 2022. ISSN 2045-2322. doi: 10.1038/s41598-022-22514-4.

E. Nakamura, T. Moritani, and A. Kanetaka. Biological age versus physical fitness age in women. 61(3):202–208, 1990. ISSN 1439-6327. doi: 10.1007/BF00357600. URL https://doi.org/10.1007/BF00357600.

Brian J. North and David A. Sinclair. The intersection between aging and cardiovascular disease. 110(8):1097–1108, 2012. doi: 10.1161/CIRCRESAHA.111.246876. URL https://doi.org/10.1161/CIRCRESAHA.111.246876. Publisher: American Heart Association.

Aäron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. abs/1807.03748, 2018. URL http://arxiv.org/abs/1807.03748.

Syed Ashiqur Rahman and Donald A. Adjeroh. Deep learning using convolutional LSTM estimates biological age from physical activity. 9(1):11425, 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-46850-0. URL https://www.nature.com/articles/s41598-019-46850-0. Publisher: Nature Publishing Group.

Luke Sagers, Luke Melas-Kyriazi, Chirag J. Patel, and Arjun K. Manrai. Prediction of chronological and biological age from laboratory data. 12(9):7626–7638, 2020. ISSN 1945-4589. doi: 10.18632/aging.102900.

Rishabh Sharma, Christoph F. Eick, and Nikolaos V. Tsekos. SM2n2: A stacked architecture for multimodal data and its application to myocardial infarction detection. In Esther Puyol Anton, Mihaela Pop, Maxime Sermesant, Victor Campello, Alain Lalande, Karim Lekadir, Avan Suinesiaputra, Oscar Camara, and Alistair Young, editors, *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, pages 342–350. Springer International Publishing, 2021. ISBN 978-3-030-68107-4.

Stephen M. Smith, Diego Vidaurre, Fidel Alfaro-Almagro, Thomas E. Nichols, and Karla L. Miller. Estimation of brain age delta from brain imaging. 200:528–539, 2019. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2019.06.017.

Jérémie Stym-Popper, Nathan Painchaud, Pierre-Yves Courand, Clément Rambour, Nicolas Thome, and Olivier Bernard. DAFTED: Decoupled asymmetric fusion of tabular and echocardiographic data for cardiac hypertension diagnosis. In *Proceedings of Machine Learning Research*, 2025. URL https://openreview.net/forum?id=ghhGImwv07&referrer=%5Bthe%20profile%20of%20Olivier%20Bernard%5D(%2Fprofile%3Fid%3D~Olivier_Bernard1).

Ye Ella Tian, Vanessa Cropley, Andrea B. Maier, Nicola T. Lautenschlager, Michael Breakspear, and Andrew Zalesky. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. 29(5):1221–1231, 2023. ISSN 1546-170X. doi: 10.1038/s41591-023-02296-6. URL https://doi.org/10.1038/s41591-023-02296-6.

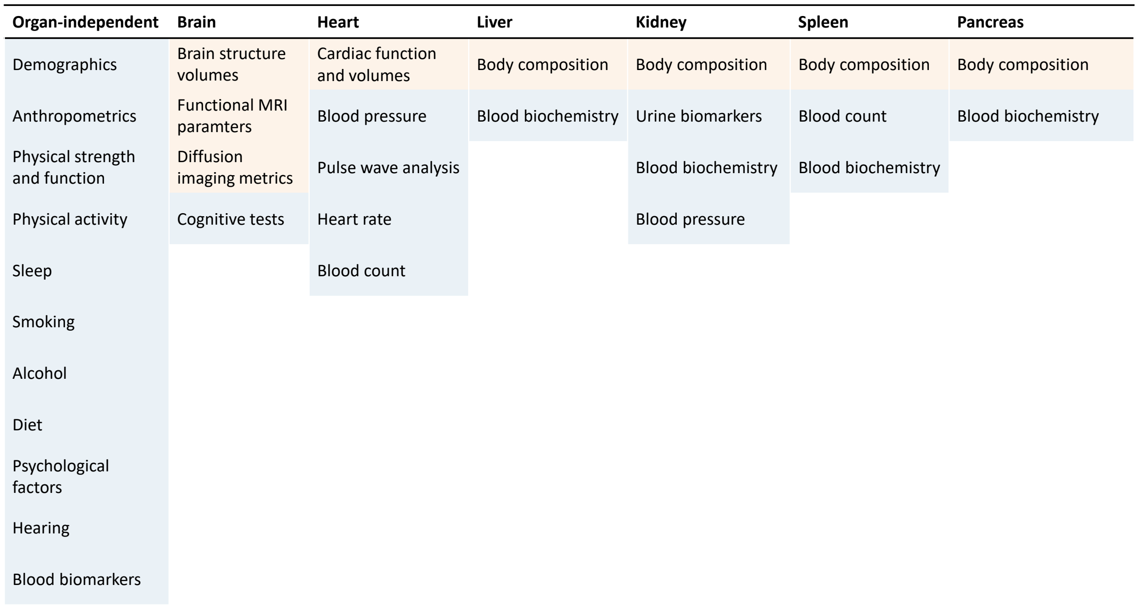
null Yanru Bai, null Xin Chen, null Changfeng Dong, null Yingxia Liu, and null Zhiguo Zhang. A comparison of multimodal biomarkers for chronic hepatitis b assessment using recursive feature elimination. 2016:2448–2451, 2016. ISSN 2694-0604. doi: 10.1109/EMBC.2016.7591225.

Hong-Yu Zhou, Yizhou Yu, Chengdi Wang, Shu Zhang, Yuanxu Gao, Jia Pan, Jun Shao, Guangming Lu, Kang Zhang, and Weimin Li. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. 7 (6):743–755, 2023. ISSN 2157-846X. doi: 10.1038/s41551-023-01045-x. URL https://doi.org/10.1038/s41551-023-01045-x.

Appendix A. Supplementary Material

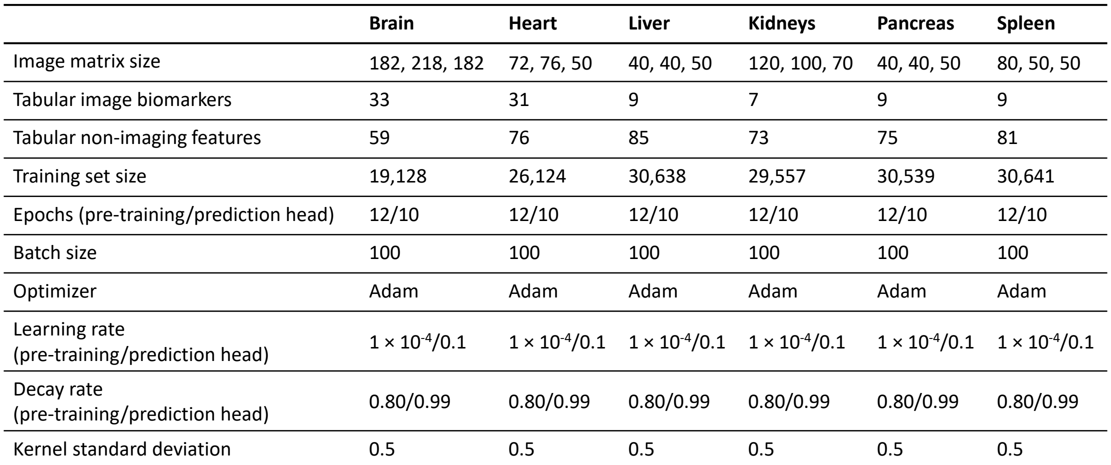
Brain	ICD-10	Fundus	ICD-10	Heart	ICD-10	Kidneys	ICD-10	Liver	ICD-10	Spleen	ICD-10	Pancreas	ICD-10
Alzheimer Disease	G30	Retinal vascular occlusions	H34	Angina Pectoris	I20	Acute renal failure	N17	Alcoholic liver disease	K70	Splenomegaly	R16.1	Acute Pancreatitis	K85
Dementia	F0*	Diabetic retinopathy	H36.0	Myocardial Infarction	I21	Chronic kidney disease	N18	Toxic liver disease	K71	Malignant neoplasm: Spleen	C26.1	Malignant neoplasm of pancreas	C25
Malignant neoplasm of brain	C71	Glaucoma	H40	Chronic Ischaemic Heart Disease	I25	Unspecified kidney failure	N19	Hepatic failure	K72	Injury of spleen (Haematoma, Laceration, Rupture of spleen)	S36	Diabeters Type 1	E10
Epilepsy	G40	Optic neuritis	H46	Cardiomyopathy	I42	Acute tubulo-interstitial nephritis	N10	Chronic hepatitis	K73	Diseases of spleen (Hyposplenism, Hypersplenism, Abscess of spleen, Cyst of spleen, Infarction of spleen)	D73	Diseases of Pancreas (Chronic Pancreatitis, Cyst/Pseudocyst of Pancreas)	K86
Mental Disorders	F0*-F9*	Other retinal disorders (Retinopathy, Degeneration of Macula and Posterior Pole, Peripheral retinal degeneration, Hereditary retinal dystrophy, Retinal haemorrhage, Separation of retinal layers)	H35	Cardiac Arrest	I46	Chronic tubulo-interstitial nephritis	N11	Fibrosis and cirrhosis of liver	K74	Diabetes Mellitus II	E11	Diabetes Mellitus II	E11
Parkinson disease	G20	Malignant neoplasm: Retina	C69.2	Atrial Fibrillation	I48	Tubulo-interstitial nephritis, not specified as acute or chronic	N12	Other inflammatory liver diseases	K75	Disorders of lipoprotein metabolism and other lipidaemias	E78	Disorders of lipoprotein metabolism and other lipidaemias	E78
Stroke	I63.9	Diabetes Mellitus II	E11	Arrhythmia	I49	Obstructive and reflux uropathy	N13	Hepatomegaly	R16.0	Metabolic syndrome	E88.81	Metabolic syndrome	E88.81
Transient cerebral ischaemic attacks	G45	Disorders of lipoprotein metabolism and other lipidaemias	E78	Congestive Heart Failure	I50.9	Drug- and heavy-metal-induced tubulo-interstitial and tubular conditions	N14	Type 2 diabetes mellitus	E11	Impaired fasting glucose	R73.1	Impaired fasting glucose	R73.1
Amytrophic Lateral Sclerosis	G12	Metabolic syndrome	E88.81	Carotid Artery Disease	I65	Other renal tubulo-interstitial diseases (Balkan nephropathy, Renal and perinephric abscess)	N15	Malignant neoplasm of liver	C22	Hyperglycemia	R73.9	Hyperglycemia	R73.9
Ataxia	R27.0	Impaired fasting glucose	R73.1	Peripheral Vascular Disease	I73.9	Cystic Kidney Disease	Q61	Diabetes Mellitus II	E11				
Bell palsy	G51	Hyperglycemia	R73.9	Hypertension	I10	Systemic Lupus Erythematosus	M32	Disorders of lipoprotein metabolism and other lipidaemias	E78				
Meningitis	G03			Hypertensive Heart Disease	I11	Glomerular disorders	N08	Metabolic syndrome	E88.81				
Guillain-Barr Syndrome	G61.0			Aortic aneurysm and dissection	I71	Haematuria	N02	Impaired fasting glucose	R73.1				
Intracranial injury	S06			Diabetes Mellitus II	E11	Cystinosis	E72.04	Hyperglycemia	R73.9				
Crushing Injury of Head	S07			Disorders of lipoprotein metabolism and other lipidaemias	E78	Calculus of Kidney	N20.0						
Traumatic amputation of part of head	S08			Metabolic syndrome	E88.81	Chronic tubulo-interstitial nephritis	N11						
Encephalitis	G04			Impaired fasting glucose	R73.1	Malignant neoplasm of kidney	C64						
Acute Spinal Cord Injury	S14.1			Hyperglycemia	R73.9	Diabetes Mellitus II	E11						
Cerebral Atherosclerosis	I67.2					Disorders of lipoprotein metabolism and other lipidaemias	E78						
Ischaemic Cerebrovascular Disease	I67.9					Metabolic syndrome	E88.81						
Diabetes Mellitus II	E11					Impaired fasting glucose	R73.1						
Disorders of lipoprotein metabolism and other lipidaemias	E78					Hyperglycemia	R73.9						
Metabolic syndrome	E88.81												
Impaired fasting glucose	R73.1												
Hyperglycemia	R73.9												

Figure 5: Excluded diseases from healthy training cohort for each organ.



Organ-independent	Brain	Heart	Liver	Kidney	Spleen	Pancreas
Demographics	Brain structure volumes	Cardiac function and volumes	Body composition	Body composition	Body composition	Body composition
Anthropometrics	Functional MRI paramters	Blood pressure	Blood biochemistry	Urine biomarkers	Blood count	Blood biochemistry
Physical strength and function	Diffusion imaging metrics	Pulse wave analysis		Blood biochemistry	Blood biochemistry	
Physical activity	Cognitive tests	Heart rate		Blood pressure		
Sleep		Blood count				
Smoking						
Alcohol						
Diet						
Psychological factors						
Hearing						
Blood biomarkers						

Figure 6: Overview of tabular features. Data fields were organized and summarized into higher-level categories. Image-derived biomarkers are marked in orange and non-imaging variables in blue.



	Brain	Heart	Liver	Kidneys	Pancreas	Spleen
Image matrix size	182, 218, 182	72, 76, 50	40, 40, 50	120, 100, 70	40, 40, 50	80, 50, 50
Tabular image biomarkers	33	31	9	7	9	9
Tabular non-imaging features	59	76	85	73	75	81
Training set size	19,128	26,124	30,638	29,557	30,539	30,641
Epochs (pre-training/prediction head)	12/10	12/10	12/10	12/10	12/10	12/10
Batch size	100	100	100	100	100	100
Optimizer	Adam	Adam	Adam	Adam	Adam	Adam
Learning rate (pre-training/prediction head)	$1 \times 10^{-4}/0.1$	$1 \times 10^{-4}/0.1$	$1 \times 10^{-4}/0.1$	$1 \times 10^{-4}/0.1$	$1 \times 10^{-4}/0.1$	$1 \times 10^{-4}/0.1$
Decay rate (pre-training/prediction head)	0.80/0.99	0.80/0.99	0.80/0.99	0.80/0.99	0.80/0.99	0.80/0.99
Kernel standard deviation	0.5	0.5	0.5	0.5	0.5	0.5

Figure 7: Training parameters of contrastive pre-training and supervised training of the age prediction head.